PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 5/222

A1

(11) International Publication Number: **WO 99/30488**

(43) International Publication Date: 17 June 1999 (17.06.99)

(21) International Application Number: PCT/IL98/00596

(22) International Filing Date: 7 December 1998 (07.12.98)

(30) Priority Data:
122498 7 December 1997 (07.12.97) IL

(71) Applicant *(for all designated States except US)*: CONTENT-WISE LTD. [IL/IL]; Building 14, 1st floor, Science Park, Nes Ziona (IL).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: HERSHTIK, Asher [IL/IL]; Modi 'Im Mobile Post, 71935 Dolev (IL).

(74) Agents: COLB, Sanford, T. et al.; Sanford T. Colb & Co., P.O. Box 2273, 76122 Rehovot (IL).

(81) Designated States: AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

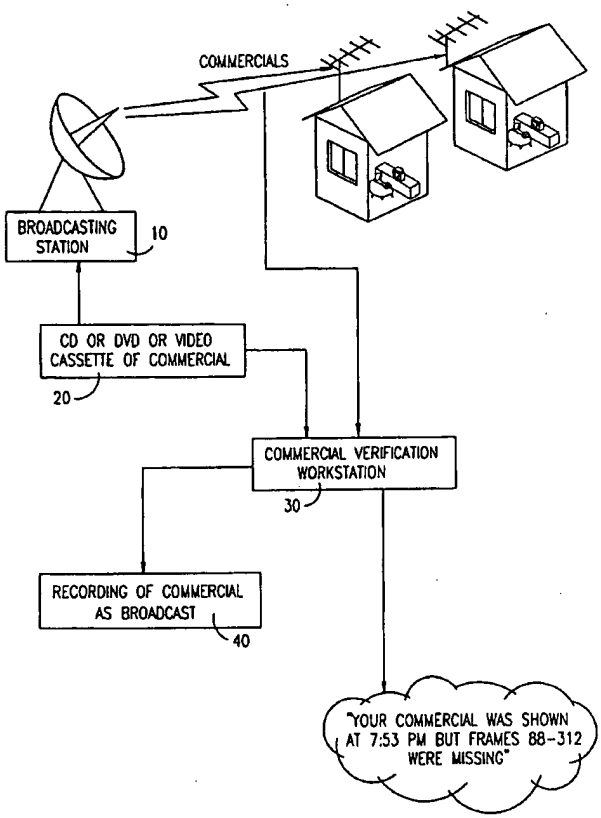
With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: APPARATUS AND METHODS FOR MANIPULATING SEQUENCES OF IMAGES

(57) Abstract

This invention discloses a video sequence viewing apparatus including an image sequence display unit (110) operative to display a sequence of images at a speed determined in accordance with a control signal, and an image sequence analyzer (100) operative to perform an analysis of the sequence of images and to generate the control signal in accordance with a result of the analysis. A watermarking method including providing an image sequence to be watermarked and performing a predetermined alteration of the length of the image sequence is also disclosed.



THIS PAGE BLANK (USPTO)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

APPARATUS AND METHODS FOR MANIPULATING SEQUENCES OF IMAGES

FIELD OF THE INVENTION

The present invention relates to apparatus and methods for manipulating sequences of images.

BACKGROUND OF THE INVENTION

Issued US Patent No. 5,790,236, entitled "Movie Processing System", inventors Asher Hershtik and Dani Rozenbaum, assignees ELOP Electronics Industries Ltd., Rehovot, Israel and Television Multilingue S.A., Geneva, Switzerland, date of issue Aug. 4, 1998, describes a movie processing system in which a plurality of versions of a movie are compared, including a movie version synchronizer and an output movie generator receiving a synchronization signal, representing the mutual synchronization of the movie versions, from the synchronizer and generating therefrom an output movie editing list.

Israel Patent Application No. 119504 describes a system and method for audio-visual content verification.

"Intro" is a known function in audio applications in which a user of a CD player can "scan" a CD by hearing a small portion of each audio segment (e.g. song) on the CD.

The disclosures of all publications mentioned in the specification and of the publications cited therein are hereby incorporated by reference.

SUMMARY OF THE INVENTION

The present invention seeks to provide improved apparatus and methods for manipulating sequences of images.

There is thus provided in accordance with a preferred embodiment of the present invention a system for capturing the signature of video frames, using only small amounts of data. The video signature technology typically captures a small amount of data characterizing each frame. The applicability of the invention includes all uses that require video identification, without the necessity of viewing.

Preferably, the system of the present invention has a PC-based platform and is operative in real-time to analyze motion pictures, video and broadcasting, inter alia.

The system of the present invention typically uses small amounts of data, to capture a signature from a stream of video frames. The signature is then matched to a continuous stream of data.

Preferably, the system of the present invention includes a matcher which synchronizes various versions of a motion picture for diverse multi-language needs including but not limited to satellite TV broadcasts, on-board film projections and DVD authoring. Another application for the system of the present invention is simplification of the restoration of damaged films by using the best footage from different versions. Yet another application is rapid adaptation of sound tracks for colorized movies.

The matcher subunit typically does not digitize video sources but rather fingerprints pictures. As a result, the matcher can process substantially any video source, such as a S-VHS video source or a 1" video source. Typically, a cassette is inserted, and a check-list is employed to choose the language to be used as a reference for matching. The user then presses PLAY and the matcher autonomously and typically without user intervention registers the fingerprint of each frame. This procedure is repeated for the next language version of the film to be checked (cassette insertion, language

selection, play). After the various versions have been fingerprinted, the versions are automatically matched, showing the differences that were detected.

The matcher preferably is operative to generate any of a variety of outputs. For example: if it is desired to broadcast multiple language versions of a film simultaneously on satellite TV, the versions must be synchronized, matcher can generate an EDL (editing list) based on the shots common to all the versions. In multi-language DVD applications, the matcher may be operative to automatically generate a branching instruction list, based on 'holes' caused by missing data in the various versions.

The system of the present invention also preferably includes a synopter for efficient viewing of video sequences. Applications include stock footage, rushes and speed-viewing of selected (typically user-selected) items of interest.

The system of the present invention also preferably includes a storyboard application which displays the first frame of every shot in an image sequence, thereby to facilitate fast-tracking of shots from rushes or stock footage. This application can operate as a search option for professional and home-use. The technology shown and described herein may be integrated into VCR's, thereby facilitating speed-searching.

For example, a user may press a first activating button and as a result, his VCR automatically adjusts search speed according to the amount of action in any given scene of a movie: slower for action-packed sequences and faster for less active moments. If the user presses a second activating button, the VCR automatically screens the first few seconds of every shot in a video, allowing the user to quickly preview the video's content.

Controlling and registering transmission of commercial spots is one of the broadcaster's most tedious

jobs. The system of the present invention preferably includes a spot shotter which monitors the off-air signal, detecting the exact moment when specific portions of any given transmission are broadcast, and automatically logging relevant information such as time of transmission and duration.

For example, the spot shotter may be "told" to detect every appearance of commercials belonging to a particular manufacturer.

Another difficult, time-consuming function for which the system of the present invention preferably is suited is automatic checking of video dubs for uniformity of content.

There is thus provided, in accordance with a preferred embodiment of the present invention, video sequence viewing apparatus including an image sequence display unit operative to display a sequence of images at a speed determined in accordance with a control signal, and an image sequence analyzer operative to perform an analysis of the sequence of images and to generate the control signal in accordance with a result of the analysis.

Further in accordance with a preferred embodiment of the present invention, the analysis of the sequence of images includes an analysis of the amount of motion in different images within the sequence and the control signal receives a value corresponding to relatively high speed for images in which there is a small amount of motion and a value corresponding to relatively low speed for images in which there is a large amount of motion.

Also provided, in accordance with another preferred embodiment of the present invention, is image sequence viewing apparatus including a shot identifier operative to perform an analysis of a sequence of images and to identify shots within the sequence of images, and

an image sequence display unit operative to sequentially display at least one initial images of each identified shot.

Further in accordance with a preferred embodiment of the present invention, the image sequence display unit is operative to display the at least one initial images of each identified shot in response to a user request.

Still further in accordance with a preferred embodiment of the present invention, the image sequence display unit is operative to display the at least one initial images of all shots sequentially until stopped by the user.

Also provided, in accordance with another preferred embodiment of the present invention, is a display system for displaying a first image sequence as aligned relative to a second, related image sequence, the system including an image sequence analyzer operative to generate a representation of a first image sequence including at least one row of pixels of each image in the first image sequence, and an aligned image sequence display unit operative to display the rows generated by the analyzer, side by side, in a single screen, wherein gaps are provided between the rows, in order to denote images which are missing, relative to the second image sequence.

Further in accordance with a preferred embodiment of the present invention, the at least one row includes at least one horizontal row of pixels and at least one vertical row of pixels.

Still further in accordance with a preferred embodiment of the present invention, the display unit is operative to display an isometric view of a stack of the images in at least one of the first and second image sequences.

Additionally in accordance with a preferred

embodiment of the present invention, the stack includes a horizontal stack.

Further in accordance with a preferred embodiment of the present invention, the analyzer also includes an image sequence aligner operative to align the first and second image sequences to one another and to provide an output denoting images which are missing from the first image sequence, relative to the second image sequence.

Additionally provided, in accordance with yet another preferred embodiment of the present invention, is a copyright monitoring system including ar image sequence comparing unit operative to conduct a comparison between an original image sequence and a suspected pirate copy of the original image sequence and to generate copyright information describing infringement of copyright of the original image sequence by the suspected pirate copy, and a copyright infringement information generator operative to generate a display of the copyright information.

Further in accordance with a preferred embodiment of the present invention, at least a portion of the comparison is conducted at the shot level.

Still further in accordance with a preferred embodiment of the present invention, at least a portion of the comparison is conducted at the frame level.

Further in accordance with a preferred embodiment of the present invention, the copyright information quantifies the infringement of copyright of the original image sequence by the suspected pirate copy.

Also provided, in accordance with yet another preferred embodiment of the present invention, is a watermarking method including providing an image sequence to be watermarked, and performing a predetermined alteration of the length of the image sequence.

Further in accordance with a preferred embod-

iment of the present invention, the performing step includes duplicating at least one predetermined image (e.g. frame or field) in the image sequence.

Still further in accordance with a preferred embodiment of the present invention, the performing step includes omitting at least one predetermined image (e.g. frame or field) from the image sequence.

Further in accordance with a preferred embodiment of the present invention, the image sequence analyzer is operative to generate aligned representations of the first and second image sequences and the display unit is operative to display the aligned representations on a single screen.

Also provided, in accordance with yet another preferred embodiment of the present invention, is a video sequence viewing method including displaying a sequence of images at a speed determined in accordance with a control signal, and performing an analysis of the sequence of images and generating the control signal in accordance with a result of the analysis.

Further provided, in accordance with yet another preferred embodiment of the present invention, is a an image sequence viewing method including performing an analysis of a sequence of images and to identify shots within the sequence of images, and sequentially displaying at least one initial images of each identified shot.

Additionally provided, in accordance with yet another preferred embodiment of the present invention, is a method for displaying a first image sequence as aligned relative to a second, related image sequence, the method including generating a representation of a first image sequence including at least one row of pixels of each image in the first image sequence, and displaying the rows generated by the analyzer, side by side, in a single screen, wherein gaps are provided between the rows, in order to denote images which are missing, rela-

tive to the second image sequence.

Further provided, in accordance with yet another preferred embodiment of the present invention, is a copyright monitoring method including conducting a comparison between an original image sequence and a suspected pirate copy of the original image sequence and to generate copyright information describing infringement of copyright of the original image sequence by the suspected pirate copy, and generating a display of the copyright information.

Still further provided, in accordance with yet another preferred embodiment of the present invention, is a watermarking system including an image sequence input device operative to input an image sequence to be watermarked, and an image sequence length alteration device operative to perform a predetermined alteration of the length of the image sequence.

BRIEF DESCRIPTION OF THE DRAWINGS AND APPENDIX

The present invention will be understood and appreciated from the following detailed description, taken in conjunction with the drawings and appendix in which:

Fig. 1 is a simplified block diagram illustration of a commercial verification system constructed and operative in accordance with a preferred embodiment of the present invention;

Fig. 2 is a simplified flowchart illustration of a preferred method of operation for the system of Fig. 1;

Fig. 3 is a simplified block diagram illustration of a system for viewing image sequences at variable speed, depending on temporally local characteristics of the image sequence such as the amount of action;

Fig. 4 is a simplified flowchart illustration

of a preferred method of operation for the system of Fig. 3;

Fig. 5 is a simplified block diagram illustration of a system for finding and displaying shots in an image sequence;

Fig. 6 is a simplified flowchart illustration of a preferred method of operation for the system of Fig. 5;

Fig. 7 is a simplified block diagram illustration of a system for displaying alignment of twc image sequences;

Fig. 8 is an isometric view of an image sequence;

Fig. 9 is an example of an isometric view of three different-language versions of the same motion picture, where gaps in the representation of a particular version indicate missing images, relative to other versions;

Fig. 10 is a simplified block diagram illustration of a copyright monitoring system constructed and operative in accordance with a preferred embodiment of the present invention;

Fig. 11 is a simplified block diagram of an electronic watermarking system constructed and operative in accordance with a preferred embodiment of the present invention; and

Appendix A is a copy of Israel Patent Application No. 119504;

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Fig. 1 is a simplified block diagram illustration of a commercial verification system constructed and operative in accordance with a preferred embodiment of the present invention. Fig. 2 is a simplified flowchart illustration of a preferred method of operation for the

system of Fig. 1. It is appreciated that the system of Figs. 1 - 2 is also useful for applications other than commercial verification, such as searching for illicit use of copyrighted sequences of images.

The apparatus of Fig. 1 includes a broadcasting system 10 which broadcasts commercials provided on a suitable receptacle 20 such as a CD or DVD or video cassette. A commercial verification workstation 30 is operative to receive broadcasts from the broadcasting system (either from the air or from a receptacle which was used to store broadcast material coming from the air) and to compare the broadcasts to an original commercial residing on the receptacle 20. The workstation attempts to identify some or all of the original commercial within the broadcasted material.

Any suitable method may be used to compare the broadcast with the original commercial. Preferably, the comparison is on the frame-level, i.e. individual frames in the broadcast, or signatures thereof, are compared to individual frames in the original commercial, or signatures thereof. Shot level comparison, in which entire shots in the broadcast are compared to entire shots in the original commercial, are typically not accurate enough. Preferred methods for comparing sequences of images, such as video images, including signature extraction and signature search (steps 60 and 70 of Fig. 2) are described in issued US Patent No. 5,790,236 and in Appendix A. Preferably, the broadcast and the original commercial are compared based only on the content of the advertisement and without requiring any special additions, e.g. without external indices, special information in vertical blanks and other special additions.

The output of the workstation 30 typically includes a recording of the commercial as broadcast and an indication of the time or times at which the commercial was broadcast, plus an indication of any incomplete-

ness in the commercial as broadcast. The output may be provided on a screen, in electronic form, as hard copy or in any other suitable format.

Figs. 1 - 2 illustrate a "cooperative" application in which the original commercial is available. It is appreciated that in some applications, in which the broadcaster and/or the advertiser are non-cooperative, the original commercial may not be available. For example, commercial monitoring of a competitor's commercials may be carried out, in which case the original commercial is, of course, not available. In these cases, a first appearance of a target commercial can be identified by a human being viewing the broadcast, and this appearance of the target commercial can then be treated as the original commercial. Alternatively, commercial monitoring can be carried out without having an original commercial, i.e. without having a model to which to compare the broadcast. For example, the system may monitor recurrence of short image sequence (i.e. image sequences which correspond in length to the known range of lengths which characterize a commercial) at time intervals which correspond to known intervals between commercial breaks.

Fig. 3 is a simplified block diagram illustration of a system for viewing image sequences at variable speed, depending on temporally local characteristics of the image sequence such as the amount of action. Fig. 4 is a simplified flowchart illustration of a preferred method of operation for the system of Fig. 3.

The apparatus of Fig. 3 includes a receptacle 90 storing an image sequence and an image sequence analyzer 100 which is typically operative to derive from each image in the image sequence a signature representing at least one characteristic of the image. For example, a "span" signature may be employed, which represents the amount of action in the image. The amount of action in an image is typically defined as the rate of change between

that image and adjacent images. Preferred methods for derivation of a "span" signature is described in issued US Patent No. 5,790,236 and in Appendix A.

The analyzer typically thresholds the signature (step 140) in order to obtain a control signal having a small number of possible values, such as 3 or 4 possible values. More generally, the control signal need not be a simple thresholded version of the signature (e.g. of the span). The control signal can have only as many values as the image sequence display unit 110 has viewing speeds. However, any suitable function may be employed to assign values to the control signal as a function of the signature. For example, the values assigned to the control signal may depend in part on second or higher order derivatives of the signature variable.

The control signal is fed to an image sequence display unit 110 such as a VCR which adjusts its speed accordingly.

Different viewing speeds can be provided by mechanical display units having motors with adjustable speed. Alternatively, if the display unit is electronic, different viewing speeds may be provided by varying the rate of display of images stored in the electronic unit.

Fig. 5 is a simplified block diagram illustration of a system for finding and displaying shots in an image sequence. Fig. 6 is a simplified flowchart illustration of a preferred method of operation for the system of Fig. 5.

The system of Fig. 5 includes a receptacle 160, such as a CD, DVD or video cassette, which stores an image sequence. An image sequence display unit, such as a VCR, is operative to display the image sequence as stored on the receptacle. The image sequence is also accessed by a shot identifier 170 which is operative, preferably on-line, to identify shots in the image sequence. Any suitable method may be used to identify the shots (step 200).

Preferred methods for identifying shots are described in issued US Patent No. 5,790,236 and in Appendix A.

The shot identifier provides a control signal, based on the locations of the shots within the image sequence, to the display unit 180. The control signal typically instructs the image sequence to display a predetermined number of frames, such as one or a few frames, at each cut, i.e. at each interface between shots. In other words, the image sequence display unit typically displays the first one or few images in each shot.

If the receptacle storing the image sequence is a physical medium such as video cassette, there is typically a time-gap between the display of the frames representing the i'th shot, and the display of frames representing the (i+1)'th shot. However, if the receptacle storing the image sequence is an electronic medium, there is typically no time-gap between the display of the frames representing subsequent shots.

It is appreciated that the image sequence display unit may display initial images for all of the shots in response to a single user command. Alternatively, the user may provide a "next shot" input each time s/he wishes to view the initial images of the next shot.

Fig. 7 is a simplified block diagram illustration of a system for displaying alignment of two image sequences. The system of Fig. 7 includes two image sequence receptacles 220 and 230, such as CDs, DVDs or video cassettes, storing two respective image sequences, such as two versions of the same motion picture. The two image sequences are aligned by an image sequence aligner 240. Image sequence aligner 240 may use any suitable image sequence aligning method to align the two sequences to one another. Preferred image sequence aligning methods are described in issued US Patent No. 5,790,236 and in Appendix A.

An isometric view generator 250 is operative to generate an isometric view of each of the image sequences. A simple isometric view of an image sequence, as illustrated in Fig. 8, may comprise an isometric view of a stack of the images in the sequence, wherein each image is regarded as a one-pixel thick rectangle, wherein all visible faces of each pixel have the color value of the pixel. It is appreciated that in the isometric view of Fig. 8, the top row of each image is visible along the top of the horizontal stack and the rightmost column of each image is visible along the side of the horizontal stack.

The isometric view generator 250 receives information regarding the alignment of the two sequences to one another from the image sequence aligner 240 and introduces gaps into the isometric view so as to illustrate the alignment. The output of the isometric view generator is typically an electronic representation 260 of an isometric view of the aligned image sequences. This representation 260 is provided to an image sequence display unit 270, such as a VCR, for display. Preferably, both aligned sequences are displayed, in isometric view, on a single screen.

Fig. 9 is an example of an isometric view of three different-language versions of the same motion picture, where gaps in the representation of a particular version indicate missing images, relative to other versions. As shown, the German version is most complete and includes no gaps, the French version has one large gap (sequence of missing frames, relative to the German version) and two smaller subsequent gaps and the English version has a total of four gaps which are not in the same locations as any of the 3 gaps of the French version.

Fig. 10 is a simplified block diagram illustration of a copyright monitoring system constructed and

operative in accordance with a preferred embodiment of the present invention. The apparatus of Fig. 10 typically includes receptacles 300 and 310, which may comprise video cassettes, DVDs, CDs and the like, which respectively store an original motion picture and a suspect pirate copy thereof. The image sequences stored in receptacles 300 and 310 are accessed by an image sequence comparison unit 320 which typically operates either at shot level or at frame level, to compare the two image sequences. Any suitable method may be employed for comparison of the two image sequences such as the methods described in issued US Patent No. 5,790,236 and in Appendix A.

The output of the image sequence comparison unit 320 typically comprises copyright monitoring information such as two aligned isometric views of the original movie and the suspect pirate copy, in which gaps denote missing frames and identical frames are placed opposite one another. Alternatively or in addition, quantitative copyright monitoring information may be provided such as the number of frames in the original movie which appear in the suspect pirate copy.

Fig. 11 is a simplified block diagram of an electronic watermarking system constructed and operative in accordance with a preferred embodiment of the present invention. According to a preferred embodiment of the present invention, image sequences such as motion pictures, news clips, commercials etc. are watermarked not by tampering in any way with any particular frame, since this tampering may impair viewing quality, but rather by either removing or adding a small number of frames from or to the image sequence. The watermark of each version or each image sequence is typically stored in an electronic databank.

In the illustrated embodiment, original and pirate copies 350 and 360 respectively of a motion pic-

ture are received by a frame-level image sequence aligner 370, in electronic form, from a video cassette (after digitization) or from a CD or DVD or other suitable image sequence receptacle. The frame-level image sequence aligner 370 is operative, according to a first embodiment of the present invention, to align the image sequence of the pirate copy to the image sequence of the original copy which preferably includes a "maximal", i.e. "union" version of the motion picture whose frames include the union of all frames in all versions of the motion picture. Any suitable method may be employed to align the two image sequences, preferably at frame level. Preferred methods for alignment of image sequences are described in issued US Patent No. 5,790,236 and in Appendix A.

Once the alignment has been determined, a watermark identifier 380 is operative to attempt to compare each of a plurality of watermarks to the aligned pirate copy. Preferably, each version of a motion picture is watermarked, including the post-production version, and each subsequent version. The "post-production version" is the motion picture as originally produced, before subsequent versions are derived therefrom. Subsequent versions are typically characterized by at least one of the following:

a. Intended distribution (airline, cable TV, cinema, etc.);

b. Language

c. Censorship (X-rated, PG-rated, R-rated, etc.)

The watermarks may be defined relative to the original copy 350. For example, "Frame #4974" is typically frame no. 4974 in image sequence 350. This is advantageous because then each suspected pirate copy need only be aligned once, to the original copy 350 (e.g. the post-production copy).

Alternatively, the frame-level image sequence aligner 370 is operative, according to a second embodi-

ment of the present invention, to align the image sequence of the pirate copy to the image sequences of each watermarked version separately, rather than aligning the pirate copy image sequence only once, to the "maximal" or "union" version of the motion picture. In this embodiment, the watermark of each version need not be defined relative to the original copy 350. For example, if every 500th field is duplicated in a PG-rated version of a motion picture, this easy rule is stored rather than computing the fields, in the maximal (complete) version, which correspond to each 500th field in the PG-rated (incomplete) version.

As shown, in the illustrated example, three watermarks are stored in this system, for each of three versions of a motion picture: post-production version, airline version, and cinema version. The airline and cinema version are typically produced from the watermarked post-production version. Typically, the watermark of the post-production version is deleted when the airline, cinema, television versions, etc., are derived from the post-production version. The post-production watermark is replaced by the watermark of the version being generated. For example, if every 500th frame is duplicated in the post-production version, whereas the watermark of the airline version calls for deletion of every 1000th frame, then the airline version is generated from the post-production version as follows:

a. the duplications of each 500th frame are removed; and

b. each 1000th frame is deleted.

As shown, in the illustrated example, the post-production watermark comprises a duplication of four specific frames. The airline version watermark comprises a duplication of one frame and removal of 3 other specific frames. The cinema version watermark comprises removal of 3 specific frames.

The watermark identifier 380 is operative to indicate the version from which the pirate copy is derived. For example, if the watermark identifier 380 finds that frames 17, 479 and 19,999 in the original copy 350 are missing in the pirate copy 360, the watermark identifier puts out a suitable output indication that the pirate copy was derived from the cinema version of a film.

It is appreciated that the software components of the present invention may, if desired, be implemented in ROM (read-only memory) form. The software components may, generally, be implemented in hardware, if desired, using conventional techniques.

It is appreciated that various features of the invention which are, for clarity, described in the contexts of separate embodiments may also be provided in combination in a single embodiment. Conversely, various features of the invention which are, for brevity, described in the context of a single embodiment may also be provided separately or in any suitable subcombination.

It will be appreciated by persons skilled in the art that the present invention is not limited to what has been particularly shown and described hereinabove. Rather, the scope of the present invention is defined only by the claims that follow:

APPENDIX A

COPY OF ISRAEL PATENT APPLICATION NO. 119504

SYSTEM AND METHOD FOR AUDIO-VISUAL CONTENT VERIFICATION

Field of the Invention

The present invention relates to audio-visual test and measurement systems and more particularly to a method and apparatus for comparing a given content stream with a reference content stream for verifying the correctness of a given data stream and for detecting various content-related problems. such as missing or distorted content. as well as badly synchronized content streams such as audio or sub-titles delayed with respect to the video stream.

"Audio-visual content" is herein defined as a stream or sequence of video, audio, graphics (sub-pictures) and other data where the semantics of the data stream is of value. The term "stream" or "sequence" is of particular importance, since it is assumed that the ordering of content elements along a time or space line constitutes part of the content.

Background of the Invention

Elementary content streams may be combined to a composite stream. Starting with a simple monophonic audio or video transmission. an application which involves two video streams (for stereoscopic display), six or eight surround audio channels and several sub-picture channels can be formed. Generally, the relative alignment of these streams is highly significant and should be verified.

In known systems. an analysis is made of video signal for detecting disturbances of that signal. such as illegal colors. An "illegal color" is one that is outside the practical limit set for a particular format. Other types of video

measurement involve injecting known signals at the source and evaluating certain properties thereof at the receiving end.

With the introduction of the serial digital interface (SDI) standard, now used as a carrier for video, audio and data, error detection schemes are designed for testing data integrity. Such a scheme has already been proposed.

The known video test and measurement systems are, however, generally not capable of detecting content-related problems, such as missing or surplus frames, program time shift, color or luminance distortions which are within the acceptable parameter range, mis-alignment of content streams such as audio or sub-pictures with respect to video, etc.

In many facilities, an observer will look at the display to detect quality problems. An experienced operator may detect and interpret a variety of problems in recording and transmission. An observer can do good rule-based or subjective evaluation of video content, however, human inspection of content is costly and unpredictable. Additionally, some content-related defects cannot be detected by an observer.

As state of the art content delivery technologies such as multi-channel Digital TV, Digital Video Disk and the Internet provide more content and interactivity, content-related problems are more likely to occur, since the path from the content sources to the end-user becomes more complicated. Additionally, the huge amounts of content generated, edited, recorded and transmitted in multiple channels and multiple distribution slots (such as video-on-demand) make human inspection almost impossible.

It is therefore a broad object of the invention to provide a computerized method and system for comparing a given content stream with a reference content stream, for verifying that the given stream is in fact the correct one and to detect various content-related defects.

In many cases, the reference stream consists of the original program material and the actual stream consists of the broadcast or played content. In other cases, the designation of one stream as the reference stream is arbitrary, for example, comparing one content stream with a backup stream. However, for convenience of description hereinafter, the terms "reference content stream" and "actual content stream" will be used, without limiting the generality of the invention.

For illustrative purposes only, the invention will be described by two applications: broadcast automation and digital versatile disc (DVD) pre-mastering. This description however, is not intended to limit the generality of the invention or its applicability to other domains.

Today's multi-channel, multi-program applications cannot be controlled manually. Including commercials and program trailers, a daily schedule may consist of hundreds of video segments, intended to play seamlessly. Such a schedule is usually implemented by an automation system. The schedule is logged into the system as some form of a table (a "play-list") describing the program's name, start time, duration and source, e.g., storage media, unique identifier, time-code of first frame.

The storage media can be a tape or a digital file. Generally, the program source material is organized in an hierarchical manner. with most of the content stored off-line. The forthcoming programs are loaded on a tape machine and sometimes. as in the case of a commercial or trailer, digitized to a disk-based server. The complex paths of the various elements of content may further increase the content mismatch probability.

An example of such an automation system is the ADC-100 from Louth Automation. ADC-100 can run up to 16 lists simultaneously, and control multiple devices including disk servers, video servers. tape machines, cart machines. VTRs. switchers. character generators and audio carts. The present invention can verify the identity and integrity of the broadcast content, providing important feedback for the automation system or facility manager.

DVD is a new generation of the compact disc format which provides increased storage capacity and performance. especially for video and multimedia applications. DVD for video is capable of storing eight audio tracks and thirty-two "sub-picture" tracks. which are used for subtitles, menus. etc. These can be used to put several selectable languages on each disc. The interactive capabilities of consumer DVD players include menus with a small set of navigation and control commands. with some functions for dynamic video stream control. such as seamless branching, which can be used for playing different "cuts" of the same video material for dramatic purposes. censorship. etc. DVD-ROM. which will be used for multi-media applications. will exhibit a higher level of interactivity.

Since DVD contains multiple content streams with many options for branching from one stream to the other or combining several streams. such as a

menu or sub-titles overlaid on a video frame, one has to verify that a given set of initial settings, followed by a specific set of navigation commands, indeed produces the correct content. This step in DVD production is known as "emulation", currently designed to be performed by an observer. The present invention also allows automation of DVD emulation.

It is important to note that in DVD, the video image is composed of the motion picture stream overlaid by sub-pictures or graphics, such as sub-titling. Although all video streams and all sub-picture bitmaps are available before emulation takes place, the composite image depends on the actual user's choices and the user's "navigation" in the content tree. It is impractical to generate all possible compositions prior to emulation and use these as the reference content. Therefore, descriptors of the actual content must be compared against appropriate descriptors of the component streams.

In both broadcast or DVD applications, it may be necessary to detect video compression artifacts. While some of these are due to the mathematical compression itself, others may arise during transmission/playback, due to buffer overflow and other reasons. A common image compression artifact is "blockiness" or the visibility of edges between image blocks. Detecting artifacts in a completely rule-based manner, such as looking for these edges, may be misleading since such edges may be present in the original, uncompressed image. An image-reference based approach in which the compressed image is compared with the original image provides a good tool for algorithm evaluation. However, in a practical situation, such an image will not be available at the receiving/playback end for real-time detection of compression artifacts. It is

therefore necessary to compare compressed material with the original material, based on concise content descriptors computed from both streams.

It is an object of the present invention to provide a content verification system in which an audio-visual program broadcast or recorded on storage media can be compared with a reference program.

The audio-visual program comprises at least one video channel, or at least one audio channel, or at least one sub-picture channel comprising sub-titles, closed-captions and any kind of auxiliary graphics information which is timed synchronously with the video or audio. While in certain applications sub-pictures are embedded in the video image sequence, in other applications they are carried by a separate stream/file.

Summary of the Invention

The present invention therefore provides a method of comparing the content obtained by broadcast or playback with a reference content, including the steps of extracting frame characteristic data streams from said reference content and from actual received or playback content, aligning said streams and comparing said streams on a frame-by-frame basis.

U.S. Patent No. 5,339,166, entitled "Motion-Dependent Image Classification for Editing Purposes," describes a system for comparing two or more versions, typically of different dubbing languages, of the same feature film. By identifying camera shot boundaries in both versions and comparing sequences of shot length, a common video version, comprising camera shots which exist in all versions, can be automatically generated. While the embodiment described in

this patent allows, in principle, the location of content differences between versions at camera shot level, frame-by-frame alignment for all frames in the respective version is not performed. Further, the differences detected are in the existence or absence of video frames as a whole. In contrast, the present invention allows frame-by-frame inspection of color properties, detection of compression artifacts, audio distortions, etc.

Furthermore, in the U.S. patent, the content of each frame is fixed and characteristic data are computed from the content. The present invention, on the other hand, addresses the on-line composition of a content stream from basic content streams, such that characteristic data are pre-computed only for these basic streams. Given the branching/navigation/editing commands, a composite reference characteristic data stream is predicted from the component characteristic data stream and then compared with the actual content stream.

Moreover, the present invention does not depend on the specific format/representation of the content sources and streams. In the same application, one stream may be analog and the other digital. Additionally, one stream may be compressed and the other may be of full bandwidth. Typically, in a broadcast environment, the input will be CCIR-601 digital video and AES digital audio. Multiple audio streams may be due to different dubbing languages, as well as stereo and surround sound channels.

Generally, the extraction of characteristic data will be done in real-time, thus saving intermediate storage and also enabling real-time error detection in a broadcasting environment. However, this is not a limitation, since the present invention can be used off-line by recording both the reference and the actual

audio-visual program. When working off-line, processing can be slower than real-time or faster, depending on the computational resources. When verifying dubs or copies of video cassettes, a faster than real-time performance may be needed, depending, of course, on the availability of a suitable analog to digital converter which can cope with fast-forward video signals.

Brief Description of the Drawings

The invention will now be described in connection with certain preferred embodiments with reference to the following illustrative figures so that it may be more fully understood.

With specific reference now to the figures in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

Fig. 1 is a block diagram of a top level flow of processing of an audio-visual content verification system;

Fig. 2 is a block diagram of a circuit for storing detected content problems;

Fig. 3 schematically illustrates an array of video sequence characteristic data;

Fig. 4 schematically illustrates an array of video frame or still image spatial characteristic data;

Fig. 5 schematically illustrates a set of regions in a video frame:

Fig. 6 schematically illustrates relative location of graphics sub-pictures with respect to the video frame;

Fig. 7 is a block diagram illustrating extraction of sub-title characteristic data;

Fig. 8 is a block diagram illustrating sub-title image sequence processing;

Fig. 9 schematically depicts a record of sub-pictures characteristic data;

Fig. 10 is a block diagram illustrating derivation of audio characteristic data;

Fig. 11 is a block diagram of a circuit for the selection of anchor frames for coarse alignment:

Fig. 12 is a block diagram of a circuit for alignment of a composite stream with the component reference streams:

Fig. 13 is a block diagram of a circuit for frame verification processing; and

Fig. 14 is a block diagram of a characteristic data design workstation.

Detailed Description of Preferred Embodiments

With reference now to the drawings, Fig. 1 shows a top level flow of processing of an audio-visual content verification system according to the present invention. Reference sub-picture stream 11. video stream 12 and audio stream 13 are stored in their respective stores 14, 15 and 16, to be eventually processed by processors 17, 18 and 19. respectively. The combination of sub-pictures with video, as well as transition/branching between program segments. is applied at characteristic data level by predictor 20. driven by navigation/playback commands 21.

Actual video stream 22 and audio stream 23 are stored in their respective stores 24 and 25, to be later processed by processors 26 and 27 respectively. The video stream 22 and the corresponding characteristic data are composed of video and sub-pictures.

Once in the characteristic data stores 28 and 29, the data streams are input to the characteristic data alignment processor 30, resulting in frame-aligned characteristic data. The alignment process also results in a program time-shift value, as well as indices or time-codes of missing or surplus frames. Once the data are frame-aligned, characteristic data are compared on a frame-by-frame basis in comparator 32, yielding a frame quality report.

Fig. 2 shows means for storing detected content problems. Recently played/received video from store 24 undergoes compression in engine 34 and is then stored in buffer 35. The recently played/received audio from store 25 is directly stored in buffer 36.. Transfer controller 37 is activated by verification reports 38 to transfer the content into hard disk storage 39, where it can be later analyzed.

Fig. 3 shows an array of video sequence characteristic data 40. The list comprises image difference measures, as well as image motion vectors. These measures may include properties of the histogram of the difference image, obtained by subtracting two adjacent images, as is known *per se*. In particular, the "span" characteristic data, defined as the difference in gray levels between a high (e.g., 85) percentile and a low (e.g., 15) percentile of said histogram, was found to be useful. Alternatively, a measure of difference of intensity histogram of two adjacent images, also by a known technique, may be used.

Motion vector fields are computed at pre-determined locations while using a block-matching motion estimation algorithm. Alternatively, a more concise representation may consist of camera motion parameters, preferably estimated from image motion vector fields.

Fig. 4 shows an array of video frame or still image spatial characteristic data. The list comprises color characteristic data 41, texture characteristic data 42 and statistics derived from image regions. Such statistics may include the mean, the variance and the median of luminance values. Useful color characteristic data include the first three moments: average, variance and skewness of color components:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} p_{ij}$$

$$\sigma_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-\mu_i)^2}$$

$$s_i = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}(p_{ij}-\mu_i)^3}$$

where p_{ij} is the value of the i-th color space component of the j-th image pixel. Color spaces of convenience may include the (R,G,B) representation or the (Y,U,V), which provide luminance characteristic data through the Y component.

Texture provides measures to describe the structural composition. as well as the distribution. of image gray-levels. Useful texture characteristic data are derived from spatial gray-level dependence matrices. These include measures such as energy, entropy and correlation.

The selection of characteristic data for a specific application of content verification is important. Texture and color data are important for matching still images. Video frame sequences with significant motion can be aligned by motion characteristic data. For more static sequences. color and texture data can facilitate the alignment process.

When computing color and texture characteristic data. the region of support. that is. the image region on which these data are computed. is significant. Using the entire image, or most of it, is preferred when robustness and reduced storage are required. On the other hand. deriving multiple characteristics at numerous. relatively small image regions has two important advantages:

1) better spatial discrimination power (like a low resolution image); and
2) when overlaid by sub-picture (graphics), those regions which do not intersect with graphics data still can be matched with corresponding characteristic data of the original video frame.

Fig. 5 shows a set of regions 42 in a video frame 43, such that color or texture characteristic data are computed for each such region. Fig. 6 illustrates the relative location of graphics sub-pictures with respect to the video frame. Number 44 represents a sub-title sub-picture and number 45 represents a menu-item sub-picture.

Figs. 7 and 8 show the extraction of sub-title characteristic data. Sub-titles or closed captions in a movie are used to bring translated dialogues to the viewer. Generally, a sub-title will occupy several dozen frames. A suitable form for sub-title characteristic data is time-code-in, time-code-out of that specific sub-title, with additional data describing the sub-title bitmap. The sub-title image sequence processor 46 analyses every video frame of the sequence to detect specific frames at which sub-title information is changed. The result is a sequence of sub-title bitmaps, with the frame interval each such bitmap occupies in a time-code-in, time-code-out representation. Characteristic data are then extracted by unit 47 from the sub-title bitmap.

Fig. 8 shows the sub-title image sequence processor 46. The video image passes through a character binarization processor 48, operative to identify pixels belonging to sub-title characters and paint them white, for example, where the background pixels are painted black. At every frame, the current frame bitmap 49 is compared, or matched, with the stored sub-title bitmap from the first instance of that bitmap. At the first mismatch event, the sub-title bitmap is reported with the corresponding time-code interval, and a new matching cycle begins.

The matching process can be implemented by a number of binary template-matching or correlation algorithms. The spatial search range of the template-matching should accommodate mis-registration of a sub-title and additionally the case of scrolling sub-titles.

The characteristic data of a single sub-title should be concise and allow for efficient matching. The sub-title bitmap, usually run-length coded, is a

suitable representation. Alternatively, one could use shape features of individual characters and a sub-title text string, using OCR software.

In addition to text, sub-pictures consist of graphics elements such as bullets, highlight or shadow rectangles, etc. Useful characteristic data are obtained by using circle and rectangle detectors. Fig. 9 shows a record 50 of sub-pictures characteristic data.

Fig. 10 shows the derivation of audio characteristic data. In analog form, the signal is digitized by the arrangement comprising an analog anti-aliasing filter 51 and an A/D converter 52 and then filtered by the pre-emphasis filter 53. Spectral analysis uses a digital filter bank 54, 54^1 . . .54^n. The filter output is squared and integrated by the power estimation unit 55, 55^1 . . .55^n. The set of characteristic data is computed for each video frame duration (40 msec for PAL, or 33.3 msec for NTSC) and stored in store 56. Window duration controls the amount of averaging or smoothing used in power computation. Typically, a 60 or 50 msec window, for an overlap of 33%, can be used.

The filter bank is a series of linear phase FIR filters, so that the group delay for all filters is zero and the output signals from the filters are synchronized in time. Each filter is specified by its center frequency and its bandwidth.

In many instances, the reference characteristic data stream is not available explicitly, but has to be derived from said source characteristic data and from playback commands such as denoted in Fig. 1. A simple case is when a program consists of consecutive multiple content segments. Each such segment is specified by a source content identifier , a beginning time-code and an ending

time-code. Said reference characteristic data stream can be constructed or predicted from the corresponding segments of source characteristic data by means of concatenation. If content verification involves computing the actual content segment insertion points, these source characteristic data segments will be padded by characteristic data margins to allow for inaccuracies in insertion.

Sometimes the transitions involve not only cuts, but also dissolves or fades. When the composite image is a linear combination of two source images, some characteristic data can be predicted based on the original source data as well as the blending values. These data include, for example, color moments computed over some region of support. In alignment and verification, the predicted values are compared against the actual values.

An important step in the verification process is the frame-by-frame alignment of the characteristic data streams. The choice of the subset of characteristic data used for alignment is important to the success of that step. Specifically, frame difference measures, such as the span described above, are well suited to alignment. A coarse-fine strategy is employed, in which anchor frames are used to solve the major time-shift between the content streams. Once that shift is known, fine frame-by-frame alignment takes place.

An anchor frame is one with an unique structure of characteristic data in its neighborhood. Fig. 11 shows the selection of anchor frames for coarse alignment. Given the frame difference data, for example, the span sequence, local variance estimation is effected in estimator 57 by means of a sliding window. Processors 58 and 59 produce a list of local variance maxima which are above a suitable threshold. A consecutive processing step in processor 60 estimates the

auto-correlation of the candidate anchor frame with its frame difference data neighborhood.

In the step of reference anchor frame selection, a further criterion may be used to increase the effectiveness of the alignment step. The anchor frames are graded by uniqueness, i.e., dissimilarity with other anchor frames, to reduce the probability of false matches in the next alignment step. Uniqueness is computed by means of cross-correlation between the anchor frame and other anchor frames. By associating the number of anchor frames with a cross-correlation value lower than a specified threshold with the specific anchor frame, those frames with highest uniqueness are selected.

Uniqueness pruning is applied only to the reference anchor frames.

Given the anchor frames of reference and actual stream, coarse alignment now begins. Each reference and actual anchor frames pair such that the cross-correlation between their respective neighborhoods is above threshold and yields a plausible alignment offset, expressed in frame count. All pairs are tested and the offsets are stored in an offset histogram array. False matches passing the cross-correlation tests will be manifested as random offset values or noise in the histogram. A nominal case of time-shifted actual content, with few or no dropped frames, will yield a single peak in the histogram. In the case of a larger number of missing or surplus frames, such as a few missing frames at each transition, the voting process described above will produce several peaks, each corresponding to a significant shift.

Having solved the time-shift between corresponding stream characteristic data intervals which are bounded by matched anchor frames, the respective intervals have to be matched. The matching process can be described as a sequence of edit operators which transform the first interval of frame characteristic data to the second interval. The sequence consists of three such operators:

1) deletion of a frame from a first stream;
2) insertion of a frame to a first stream: and
3) replacement of a frame from a first stream with a frame from a second stream.

Having associated a cost with each of these operations, the fine frame alignment problem has now been transformed to finding a minimum cost sequence of operators which implements the transformation. If m is the length of the first interval and n is the length of the second interval in frames, then the matching problem can be solved in space and time proportional to $(m*n)$. All that remains is to set the respective costs. Deletion and insertion can be assigned a fixed cost each, based on *a-priori* information on the probability of dropped or surplus frames. Replacement is a distance measure on the characteristic data vector, such as weighted Euclidean distance.

Fig. 12 shows the alignment of a composite stream with the component reference streams by means of a processor 61 and geometric filter 62. In a simple case, sub-title graphics of the language of choice are combined with the video frame sequence. The location of sub-titles in the video frame can be specified either manually, in the characteristic data design workstation as described below, or can be automatically computed, based on analysis of the sub-

title sub-picture stream. For that simple case. video frame verification is done in the image region free from sub-titles. Additionally, sub-title picture verification is done in the sub-title image region.

A more difficult case is when graphics are overlaid on the video frame, such as in the case of displaying a menu in a DVD player. The location of menu bullets and text may be. for example. as illustrated in Fig. 6. For that specific case. it is assumed that the graphics stream has been pre-processed to extract the graphics regions of support. in the form of bounding rectangles for text lines and graphics primitives. These regions are stored as auxiliary characteristic data. By comparing graphics stream characteristic data with composite video frame stream graphics characteristic data in the respective graphics regions, the streams can be aligned. Once aligned. the composite frame graphics regions are known to be those of the corresponding graphics stream. Then. based on these regions, only color and texture actual frame characteristic data which are not occluded by overlay graphics [see Fig. 6] are compared with the respective reference data.

Fig. 13 depicts the frame verification processes performed by the frame characteristic data comparator 32 (Fig. 1), which start from aligned characteristic data streams. It is important to note that the characteristic data alignment processor 30 detects a variety of content problems. Failure in alignment may be due to the fact that a wrong content stream is playing, or the content stream is severely time-shifted. or the stream is distorted beyond recognition. A successful alignment yields the indices of missing or surplus frames. Once aligned. each actual content frame is compared with the corresponding reference frame, based on the characteristic data.

Then for the remaining data, frame-by-frame comparison can take place in processors 63, 64 and 65 and comparators 66 and 67. The distance between characteristic data of corresponding frames detects quality problems such as luminance or color change, as well as audio distortions. By comparing graphics characteristic data, errors in sub-picture content and overlay may be detected. Also, by comparing characteristic data sensitive to compression artifacts, such artifacts can be detected.

The comparison process requires the notions of distance and threshold. For vector characteristic data such as color, luminance and audio, a vector distance measure is used, such as the Mahalonobis distance:

$$D = (X^r - X^a)^T C^{-1} (X^r - X^a)$$

where X^r, X^a are the reference and actual characteristic data vectors. C is the co-variance matrix which models pairwise relationships among the individual characteristic data. The proper threshold may be computed at a training phase, using the characteristic data design workstation described hereinafter with reference to Fig. 14.

Comparator 68 compares blockiness characteristic data derived from the reference and actual video frames, respectively. Such data may include power estimates of a filter designed to enhance an edge grid structure, such as, for example, the grid spacing equals the compression block size, which is usually 8 or 16. By comparing these estimates with the reference value, an increase in blockiness may be detected. As described above, absolute blockiness may be misleading, since it may originate from the original frame texture.

Comparison of sub-pictures can be done at bitmap level, at the exclusive OR of the corresponding bitmaps, by computing the distance between corresponding shape characteristic data vectors, or by comparing recognized sub-title text strings, where applicable.

The term "frame-by-frame," which is used in conjunction with the comparison process, relates to the fact that once the content streams are aligned, inspection of every frame with the corresponding frame can be done. Clearly, comparison may include all frames or a sub-set of the frames.

The efficiency, robustness and content verification could be enhanced by using features that have greater discriminating power over the full reference content. By designing a software-configurable characteristic data set, the actual data of the full set which is implemented will be enabled.

Fig. 14 shows a characteristic data design workstation 69. The characteristic data acquisition part of the work-station replicates the reference content processing front-end of Fig. 1. In addition, workstation 69 has access, by network 70, to the actual content data and not just to the characteristic data, for display at 71 and further analysis at 72.

The development of the specific content verification application is conducted using an arrangement of a combination of manual, semi-automatic and automatic processes. For example, the user may specify the sub-titling type-face and its location in the video frame. Additionally, the user may select several representative content segments and the system then extracts a full characteristic data set, possibly in multiple passes or slower than real-time, ranking their

discriminating power over the sample reference content and retaining their best features.

It will be evident to those skilled in the art that the invention is not limited to the details of the foregoing illustrated embodiments, and that the present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are, therefore, intended to be embraced therein.

The method of the invention may further comprise the step of computing actual characteristic data from at least part of the actual broadcast or playback content streams. It may also comprise the step of computing reference characteristic data from at least part of said reference content streams.

Said reference characteristic data may be derived from video frame sequences, still images, audio and graphics, and said actual characteristic data may be derived from a video sequence and an audio channel. Also, said video image sequence characteristic data may include an image motion vector field, or data derived from an image difference signal, and said video frame or still image characteristic data may include luminance statistics in pre-defined regions of said frame or image.

Preferably, said video frame or still image characteristic data also include texture characteristic data and/or colour data, said colour characteristic data include colour moments, said video frame or still image characteristic data also include a low resolution or highly compressed version of the original image, said audio characteristic data include audio signal parameters, estimated at a window size which is comparable with video frame duration, said graphics characteristic data exhibit printed text, and said graphics characteristic data also exhibit common graphics elements, including bullets and highlighted rectangles

In the method of the invention, said step of predicting may include generating a characteristic data stream from source streams and navigation commands or play-lists, branching from one source stream to another source stream. Said step of predicting may also include generating a characteristic data stream from source streams and transition commands such as cut, dissolve, fade to/from black, or said step may include computing characteristic data of graphics sub-pictures overlay on a video image sequence or still.

The evaluation of the information content of a certain frame may be based on the temporal variation of characteristic data in said frame and in its adjacent frames.

The method may further comprise grading the information content of all frames in a sequence, denoting frames with locally maximal information content as anchor frames.

The method may still further comprise evaluating the similarity between two anchor points, based on a measure of temporal correlation between the respective sets of neighbouring characteristic data. Alternatively, the method may further comprise evaluating the similarity between all pairs of anchor frames, such that, for each pair, one frame is from the reference data and the other is from the actual data.

The method may further comprise reporting said alignment results, including the time shift between the designed and actual content broadcast-playback, as well as an indication of missing or surplus frames. The step of comparing may comprise first aligning the graphics of said composite frame sequence with said reference graphics streams, and the step of aligning may facilitate computing the location of all overlaid graphics in said composite frame sequence. The step of computing may facilitate filtering out colour and texture actual frame characteristic data which are occluded by said overlay graphics.

The method may further comprise comparing characteristic data of aligned frames to indicate quality or content problems, and said problems may be selected from the group comprising luminance or colour shifts, compression artifacts, audio artifacts, and audio or sub-pictures mismatch or mis-alignment.

CLAIMS.

1. A method for video content verification, operative to compare and verify the content of a first audio-visual stream with the content of a second audio-visual stream, the method comprising the steps of:

extracting characteristic data from a first audio-visual stream;

extracting characteristic data from a second audio-visual stream; and

comparing the extracted characteristic data from said first and second audio-visual streams.

2. A method as claimed in claim 1, wherein the step of comparison comprises:

aligning said first and second audio-visual streams on a frame-by-frame basis; and

performing a frame-by-frame comparison of said aligned streams of frames.

3. A method as claimed in claim 1 or claim 2, wherein said first and second streams are selected from the group comprising the elementary content streams, including video image sequence, audio channel, and sub-picture streams.

4. A method as claimed in any one of claims 1 to 3, wherein said comparison of first and second streams yields at least one parameter, including time-shift between the desired and the actual timing of said second stream; list of missing frames in said second stream; list of surplus frames in said second stream; sub-title content error; graphics content error, colour distortion, and luminance shift.

5. A method for video content verification, operative to compare and verify the content of a first audio-visual stream with the content of a second audio-visual stream, wherein said second audio-visual content stream is defined by at least one source content stream and a set of editing instructions, the method comprising the steps of:

extracting characteristic data from said first audio-visual stream;

extracting characteristic data from said source content stream, and

computing characteristic data of said second content-stream. based on characteristic data of said source content stream and on said editing instructions.

6. A method as claimed in claim 5. wherein said instructions are in the form of an Edit Decision List or Digital Video Disk branching instructions.

7. A method as claimed in any one of claims 1 to 6. wherein said first or second stream is a reference content stream.

8. A method as claimed in any one of claims 1 to 6. wherein said first and/or second streams are actual broadcast or playback content streams.

9 A method as claimed in claim 7. further comprising the step of predicting the reference characteristic data stream from said reference characteristic data and from playback instructions.

10. A method as claimed in any one of claims 1 to 9. wherein said characteristic data extraction is optionally augmented by user input facilitating the extraction/relative weighting of said data.

11. A method as claimed in claim 7. further comprising aligning the reference characteristic data stream with the actual characteristic data stream. on a frame-by-frame basis. and evaluating the information content of a certain frame.

12. A method as claimed in claim 11. further comprising computing the frame-index offset between the reference and actual frames. based on the most likely offsets derived from evaluation of the similarity between all anchor frames.

13. A method as claimed in claim 11. further comprising matching the reference frame sequence with the actual frame sequence. based on an identified frame-index offset. and

further comprising the step of designating an actual frame as a surplus frame. or assigning to it a unique reference frame.

14. A method as claimed in any one of claims 1 to 13. further comparing a composite video frame sequence including graphics overlaid on a video frame sequence, with component reference streams consisting of the original video frame sequence as well as the graphics streams.

15. A system for audio-visual content verification, operative to compare and verify the content of a first audio-visual data stream with the content of a second audio-visual data stream. the system comprising:

means for extracting characteristic data from a first audio-visual data stream;

means for extracting characteristic data from a second audio-visual data stream; and

means for comparing characteristic data of said first and second audio-visual data streams.

16. A system as claimed in claim 15, wherein said comparison means comprises:

means for aligning said audio-visual data streams on a frame-by-frame basis: and

means for frame-by-frame comparison of said aligned data streams.

17 A system as claimed in claim 15 or claim 16, wherein said first and second data streams are selected from the group comprising video image sequence. audio channel, and sub-picture data streams.

18. A system as claimed in any one of claims 15 to 17, wherein said means for comparison of said reference data streams yields at least one of the parameters including time-shift between the desired and the actual timing of said second data stream: list of missing frames in said second data stream; list of surplus frames in said second data stream: sub-title content error; graphics content error; colour distortion. and luminance shift.

SYSTEM AND METHOD FOR AUDIO-VISUAL CONTENT VERIFICATION

ABSTRACT

The invention provides a method for video content verification, operative to compare and verify the content of a first audio-visual stream with the content of a second audio-visual stream, comprising the steps of extracting characteristic data from a first audio-visual stream, extracting characteristic data from a second audio-visual stream, and comparing the extracted characteristic data from the first and second audio-visual streams. The invention also provides a system for carrying out the method.

19. A system for audio-visual content verification, operative to compare and verify the content of a first audio-visual data stream with the content of a second audio-visual data stream, wherein said second audio-visual data stream is defined by at least one source content data stream and a set of editing instructions, the system comprising:

means for extracting characteristic data from said first audio-visual data stream;

means for extracting characteristic data from said source content data stream; and

means for computing characteristic data of said second content data stream, based on characteristic data of said source content data stream and said editing instructions.

20. A system as claimed in claim 19, wherein said editing instructions are in the form of an Edit Decision List or Digital Video Disk branching instructions.

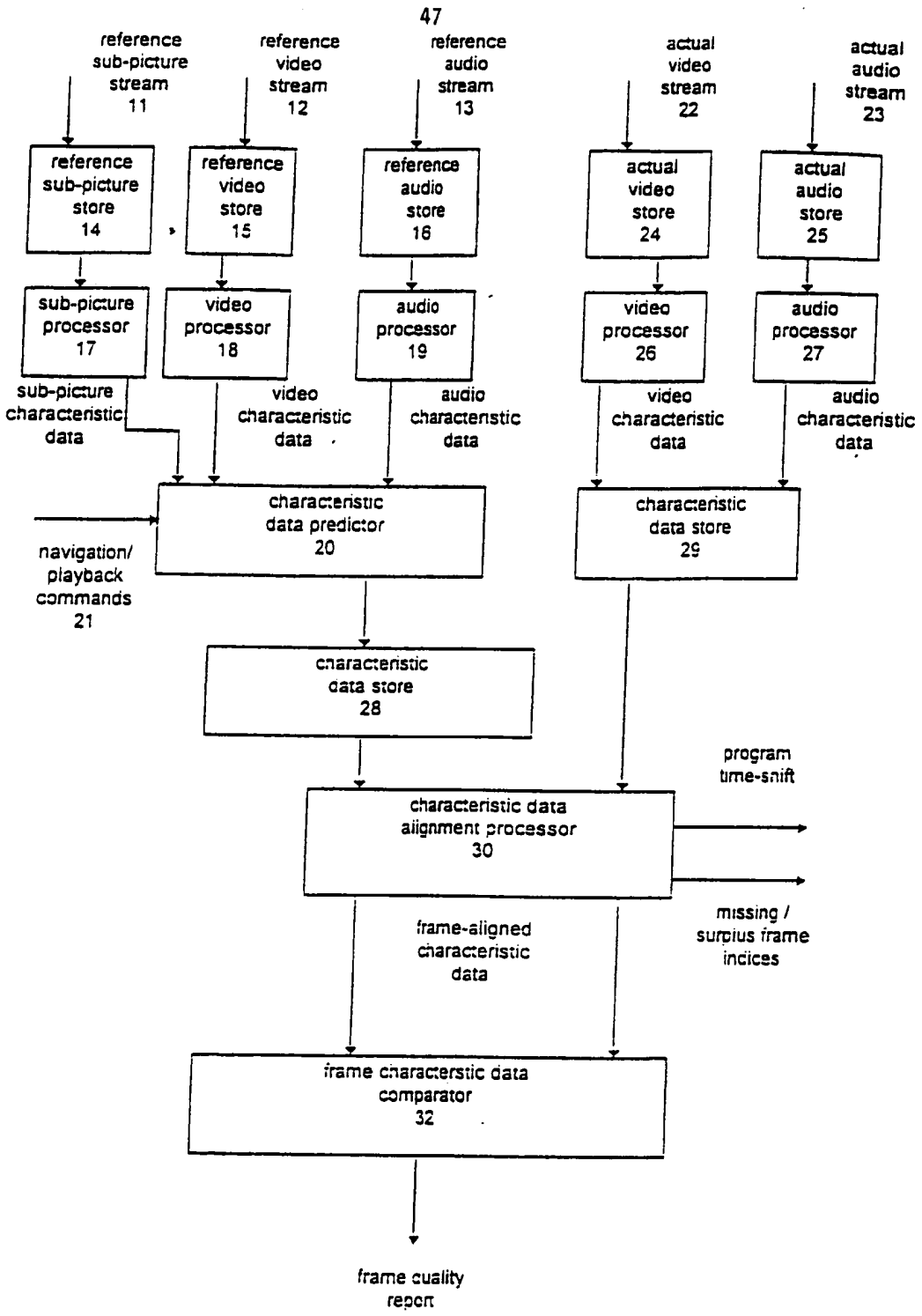

47
reference sub-picture stream 11
reference video stream 12
reference audio stream 13
actual video stream 22
actual audio stream 23
reference sub-picture store 14
reference video store 15
reference audio store 16
actual video store 24
actual audio store 25
sub-picture processor 17
video processor 18
audio processor 19
video processor 26
audio processor 27
sub-picture characteristic data
video characteristic data
audio characteristic data
video characteristic data
audio characteristic data
characteristic data predictor 20
characteristic data store 29
navigation/ playback commands 21
characteristic data store 28
program time-shift
characteristic data alignment processor 30
missing / surplus frame indices
frame-aligned characteristic data
frame characterstic data comparator 32
frame quality report

Fig. 1

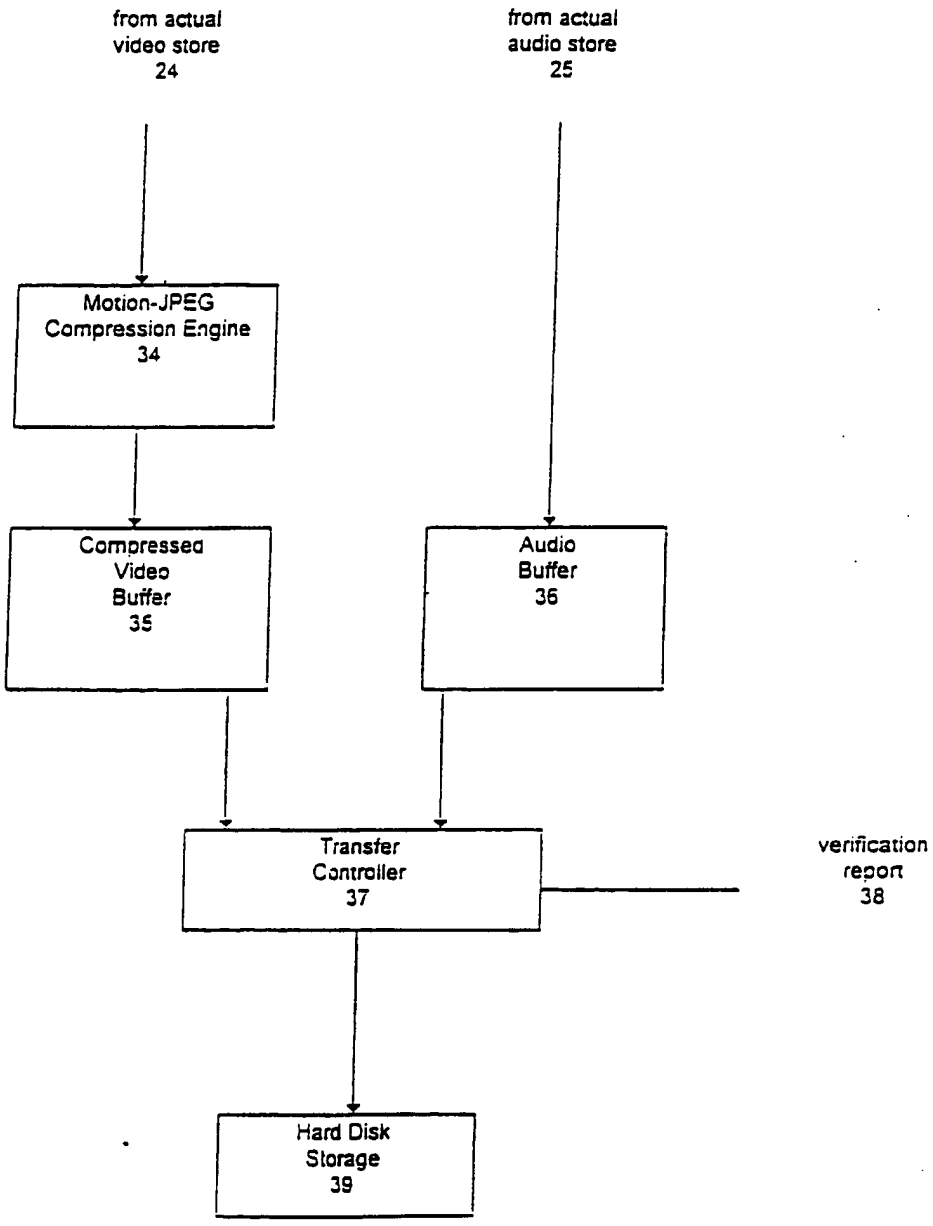

from actual
video store
24
from actual
audio store
25
Motion-JPEG
Compression Engine
34
Compressed
Video
Buffer
35
Audio
Buffer
36
Transfer
Controller
37
verification
report
38
Hard Disk
Storage
39

Fig. 2

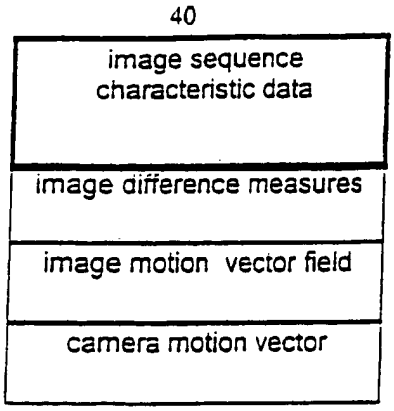


Fig. 3

41 color characteristic data	42 texture characteristic data
average	energy
variance	entropy
skewness	correlation

Fig. 4

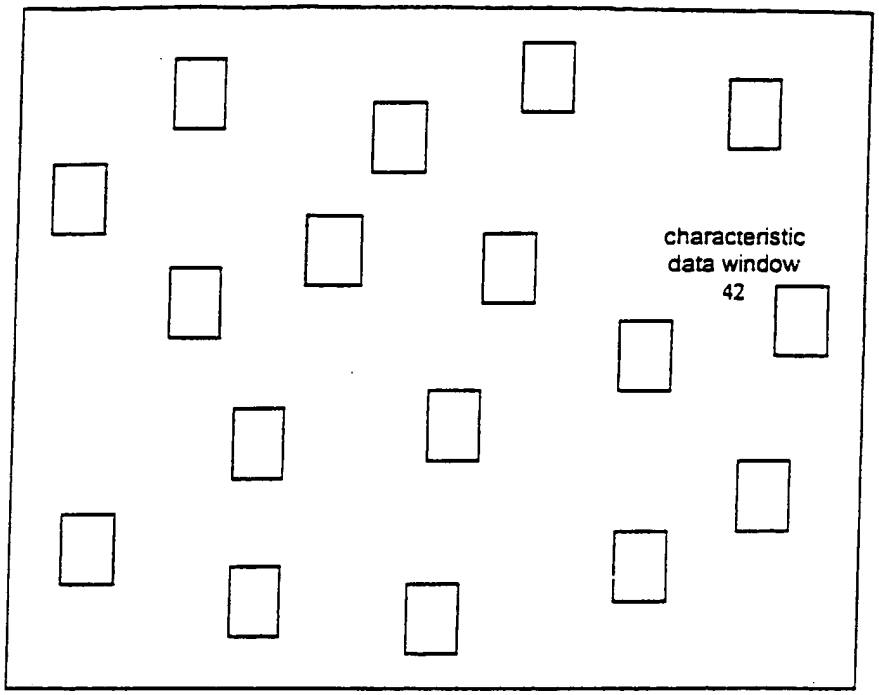

video frame
43
characteristic
data window
42

Fig. 5

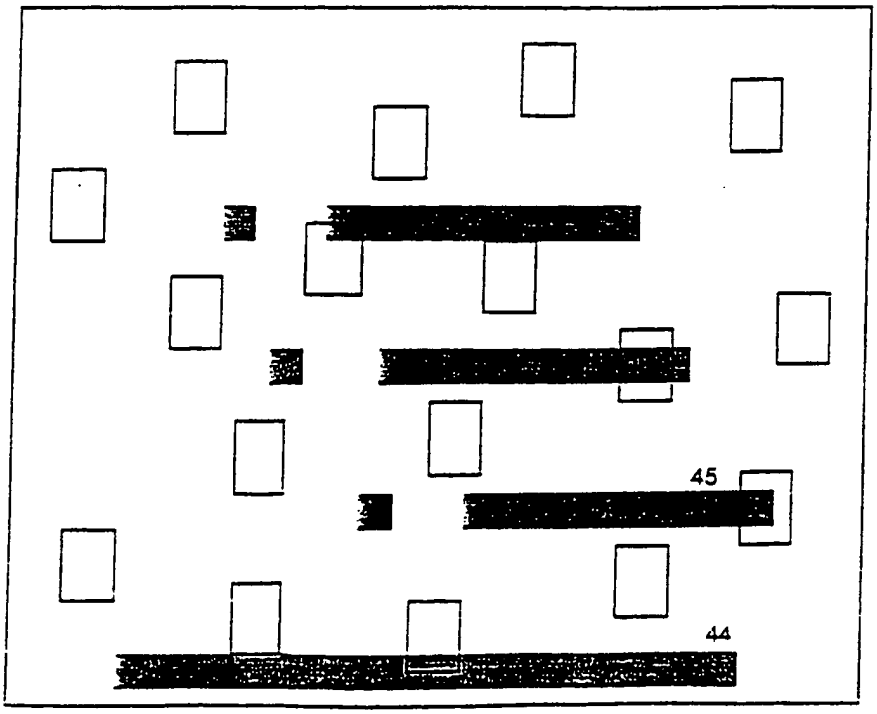

video frame
43
45
44

Fig. 6

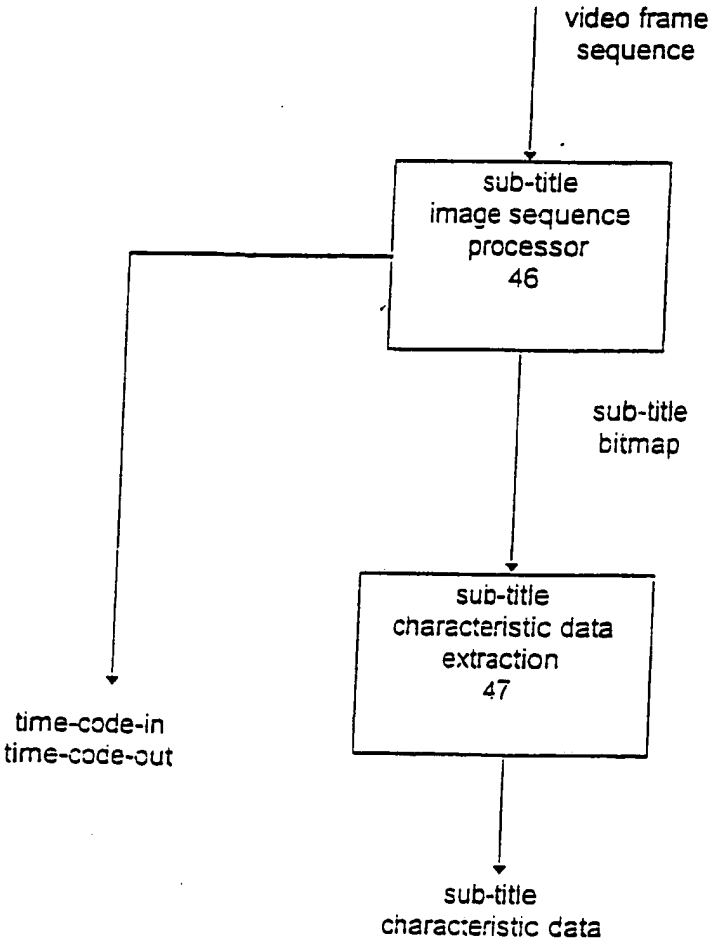

video frame sequence
sub-title image sequence processor 46
sub-title bitmap
sub-title characteristic data extraction 47
time-code-in time-code-out
sub-title characteristic data

Fig. 7

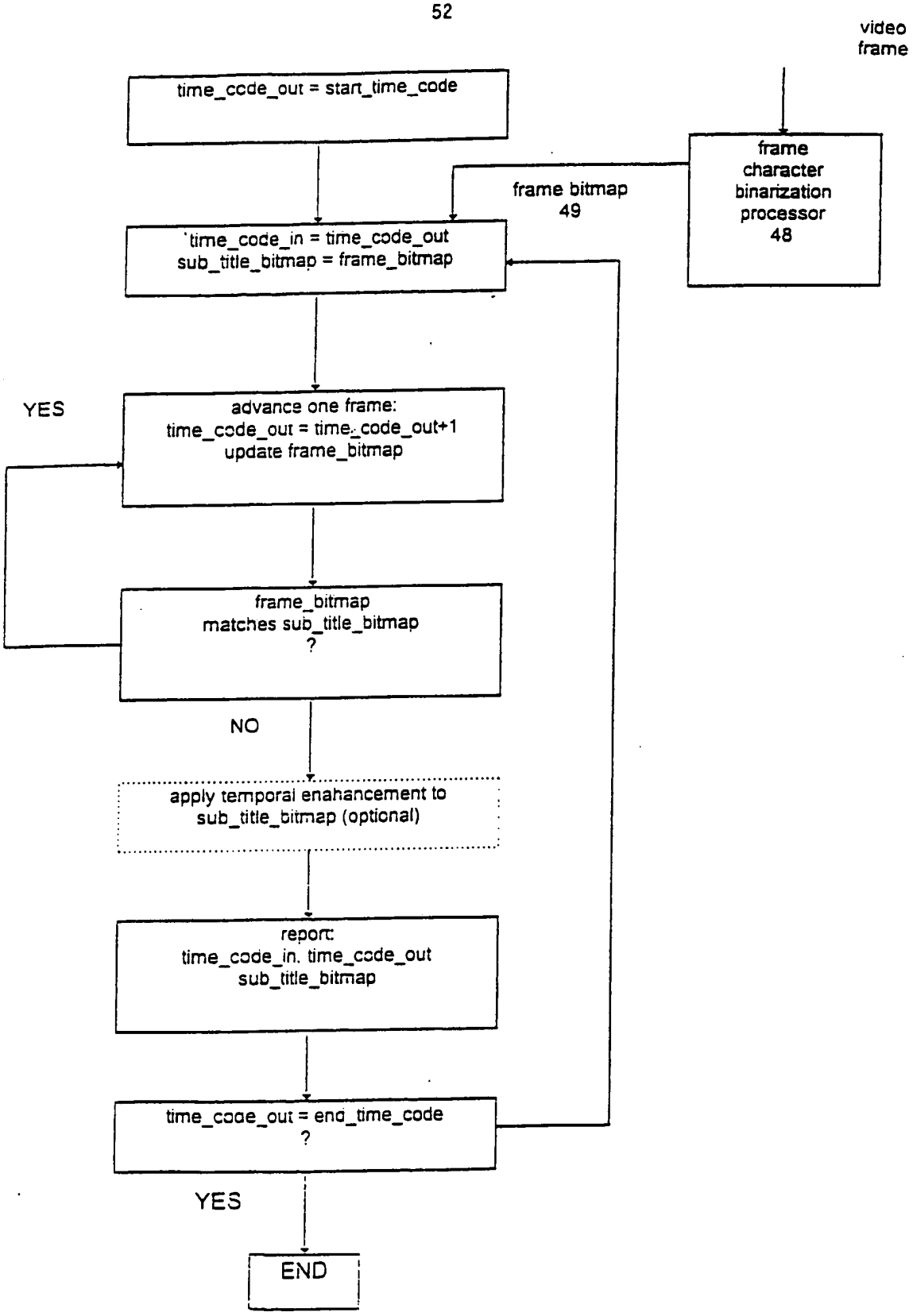

video frame
time_code_out = start_time_code
frame character binarization processor 48
frame bitmap 49
time_code_in = time_code_out
sub_title_bitmap = frame_bitmap
YES
advance one frame:
time_code_out = time_code_out+1
update frame_bitmap
frame_bitmap matches sub_title_bitmap ?
NO
apply temporal enahancement to sub_title_bitmap (optional)
report:
time_code_in. time_code_out
sub_title_bitmap
time_code_out = end_time_code ?
YES
END

Fig. 8

50

sub-title characteristic data	graphics characteristic data
sub-title bounding rectangles	highlight rectangles
sub-title bitmaps	bullets center coordinates
sub-title shape data	
sub-title text string	

Fig. 9

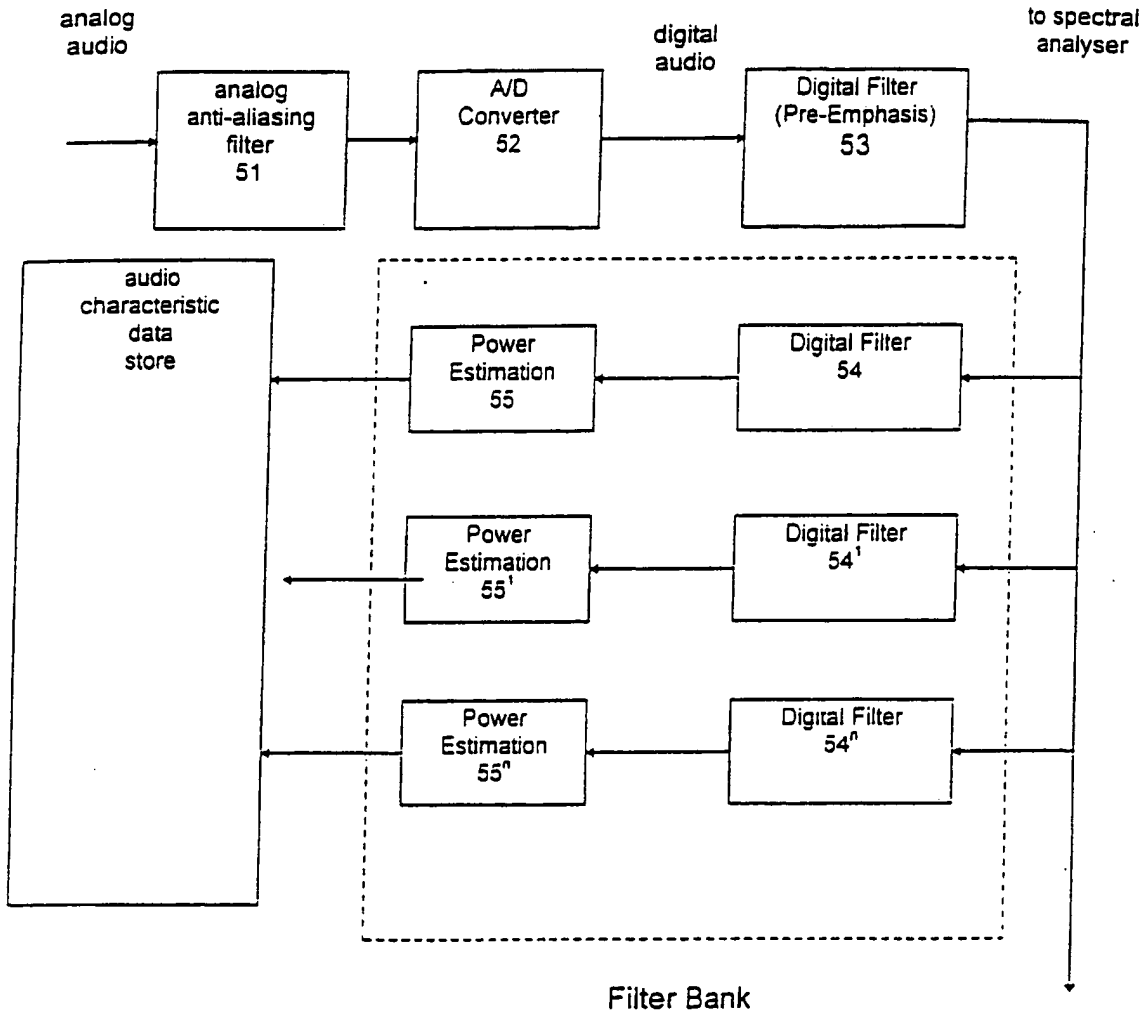

analog audio
analog anti-aliasing filter 51
A/D Converter 52
digital audio
Digital Filter (Pre-Emphasis) 53
to spectral analyser
audio characteristic data store
Power Estimation 55
Digital Filter 54
Power Estimation 55'
Digital Filter 54'
Power Estimation 55n
Digital Filter 54n
Filter Bank

Fig. 10

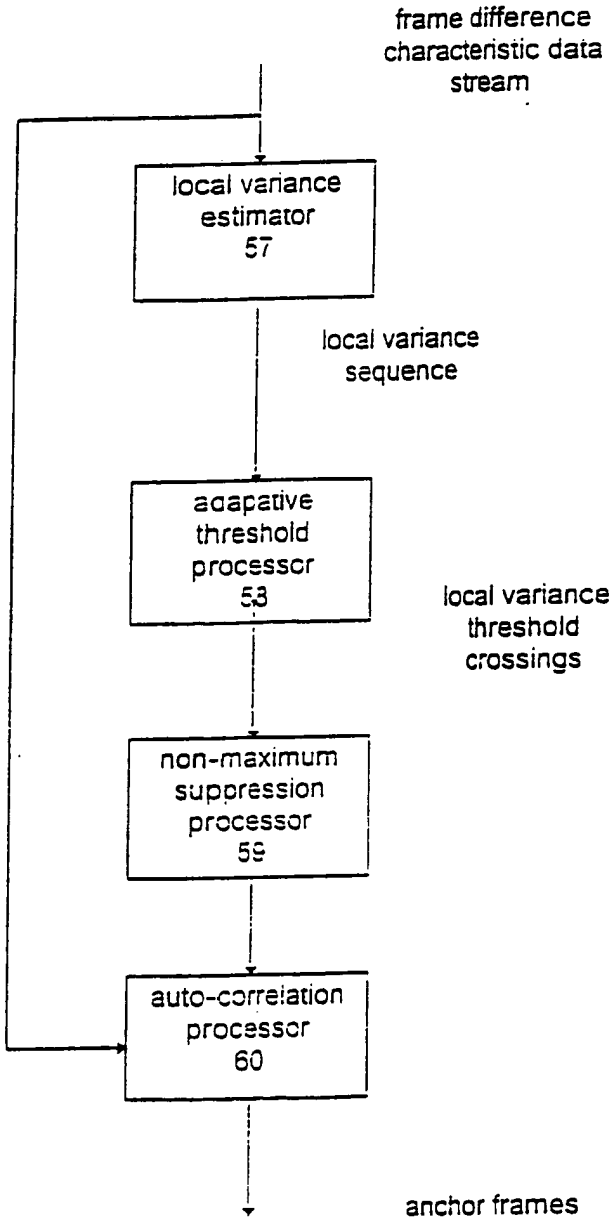

frame difference characteristic data stream
local variance estimator 57
local variance sequence
adaptive threshold processor 58
local variance threshold crossings
non-maximum suppression processor 59
auto-correlation processor 60
anchor frames

Fig. 11

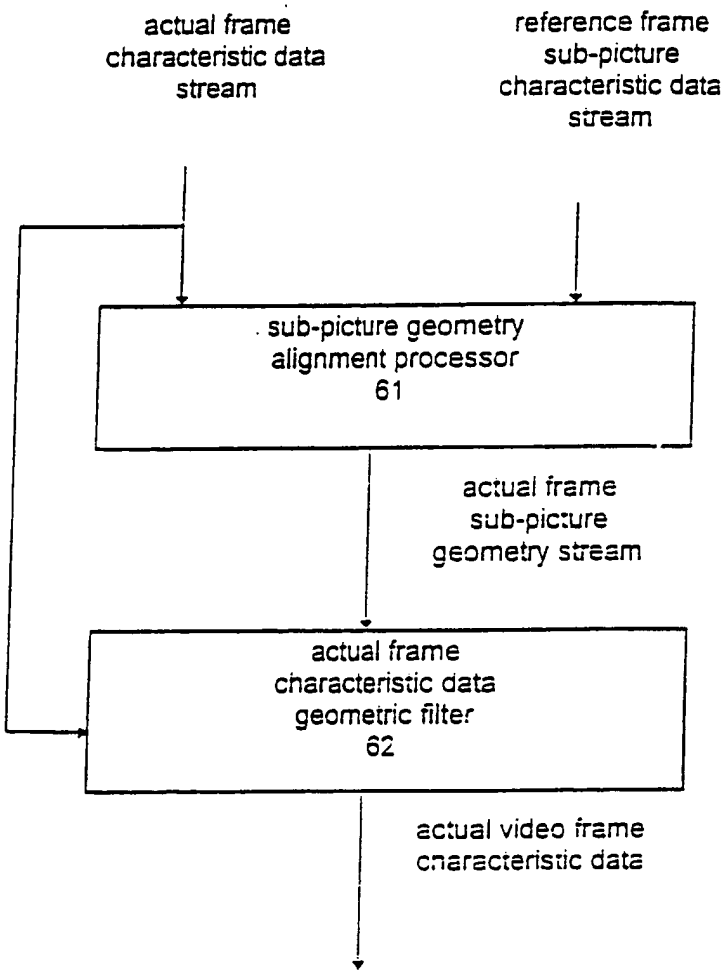

actual frame characteristic data stream
reference frame sub-picture characteristic data stream
sub-picture geometry alignment processor 61
actual frame sub-picture geometry stream
actual frame characteristic data geometric filter 62
actual video frame characteristic data

Fig. 12

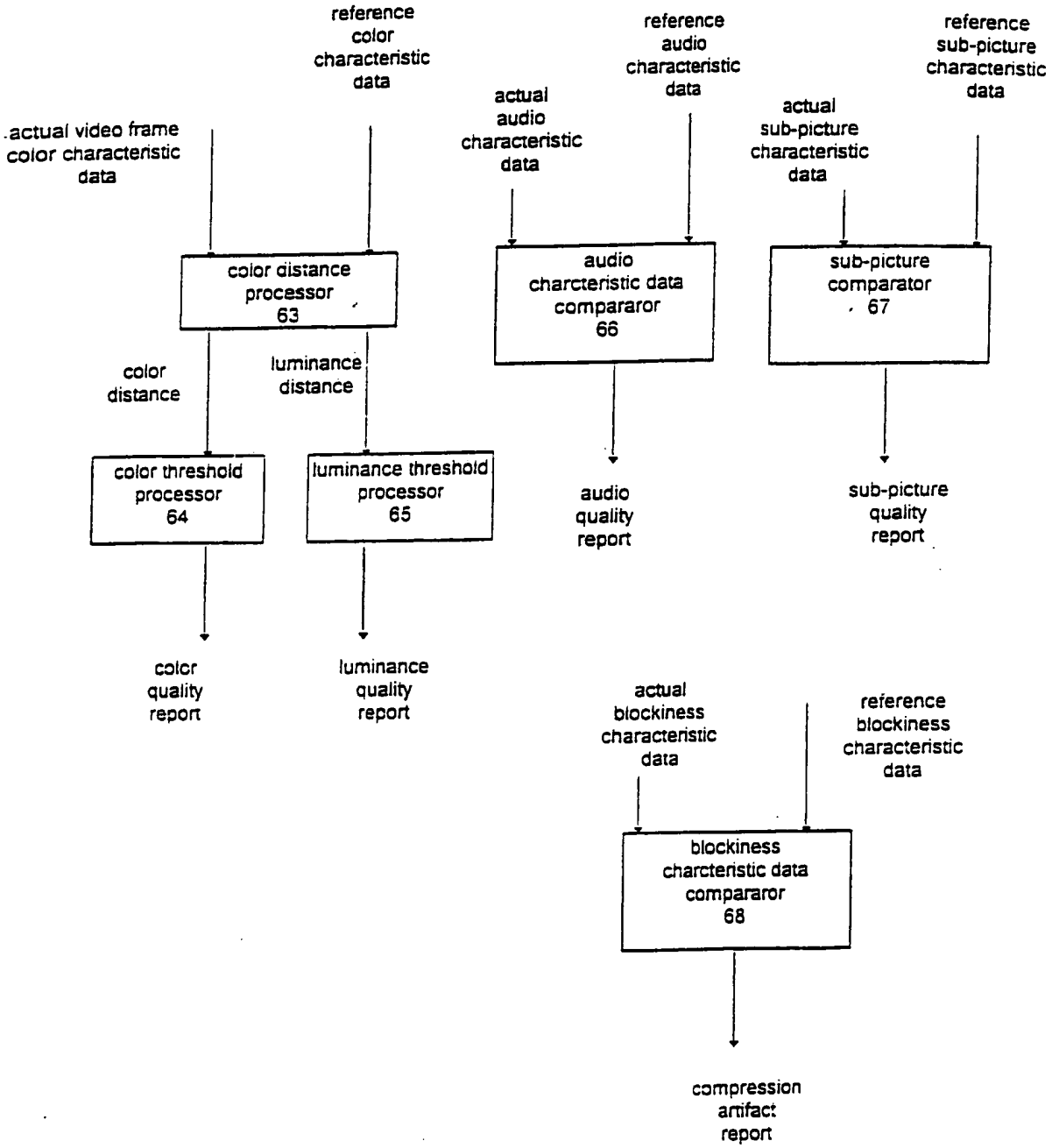

reference color characteristic data
actual video frame color characteristic data
color distance processor 63
color distance
luminance distance
color threshold processor 64
luminance threshold processor 65
color quality report
luminance quality report
reference audio characteristic data
actual audio characteristic data
audio charcteristic data compararor 66
audio quality report
reference sub-picture characteristic data
actual sub-picture characteristic data
sub-picture comparator 67
sub-picture quality report
actual blockiness characteristic data
reference blockiness characteristic data
blockiness charcteristic data compararor 68
compression artifact report

Fig. 13

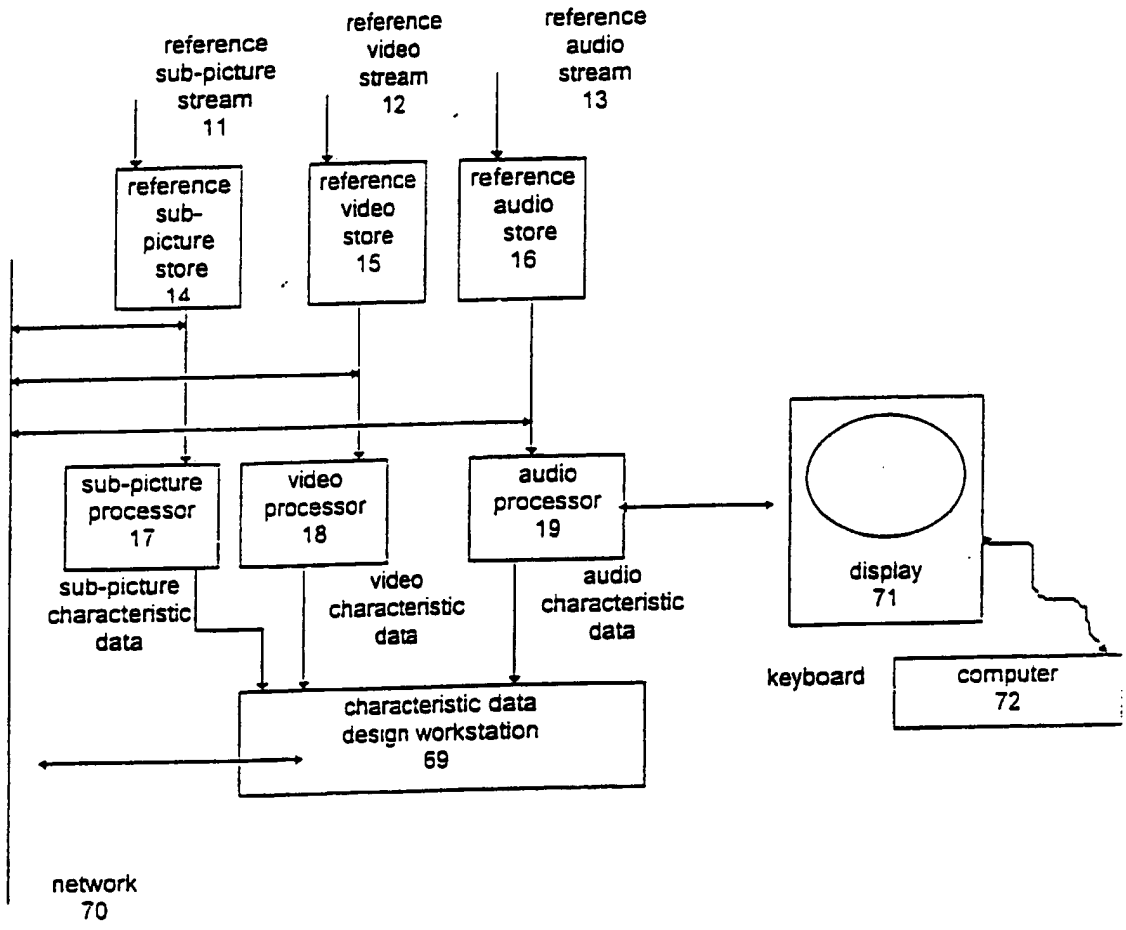

reference sub-picture stream 11
reference video stream 12
reference audio stream 13
reference sub-picture store 14
reference video store 15
reference audio store 16
sub-picture processor 17
video processor 18
audio processor 19
display 71
sub-picture characteristic data
video characteristic data
audio characteristic data
keyboard
computer 72
characteristic data design workstation 69
network 70

Fig. 14

CLAIMS

1. Video sequence viewing apparatus comprising:

an image sequence display unit operative to display a sequence of images at a speed determined in accordance with a control signal; and

an image sequence analyzer operative to perform an analysis of the sequence of images and to generate the control signal in accordance with a result of the analysis.

2. Apparatus according to claim 1 wherein the speed comprises a variable speed and the control signal has more than one values.

3. Apparatus according to claim 1 or claim 2 wherein the analysis of the sequence of images comprises an analysis of the amount of motion in different images within said sequence and said control signal receives a value corresponding to relatively high speed for images in which there is a small amount of motion and a value corresponding to relatively low speed for images in which there is a large amount of motion.

4. Image sequence viewing apparatus comprising:

a shot identifier operative to perform an analysis of a sequence of images and to identify shots within the sequence of images; and

an image sequence display unit operative to sequentially display at least one initial images of each identified shot.

5. Apparatus according to claim 4 wherein the image sequence display unit is operative to display the at least one initial images of each identified shot in response to a user request.

6. Apparatus according to claim 4 or claim 5 wherein the image sequence display unit is operative to display the at least one initial images of all shots sequentially until stopped by the user.

7. A display system for displaying a first image sequence as aligned relative to a second, related image sequence, the system comprising:

an image sequence analyzer operative to generate a representation of a first image sequence including at least one row of pixels of each image in the first image sequence; and

an aligned image sequence display unit operative to display the rows generated by the analyzer, side by side, in a single screen, wherein gaps are provided between the rows, in order to denote images which are missing, relative to the second image sequence.

8. A system according to claim 7 wherein the at least one row comprises at least one horizontal row of pixels and at least one vertical row of pixels.

9. A system according to claim 7 wherein the display unit is operative to display an isometric view of a stack of the images in at least one of the first and second image sequences.

10. A system according to claim 9 wherein the stack comprises a horizontal stack.

11. A system according to claim 7 and wherein the analyzer also comprises an image sequence aligner operative to align the first and second image sequences to one another and to provide an output denoting images which are missing from the first image sequence, relative to the second image sequence.

12. A copyright monitoring system comprising:

an image sequence comparing unit operative to conduct a comparison between an original image sequence and a suspected pirate copy of the original image sequence and to generate copyright information describing infringement of copyright of the original image sequence by the suspected pirate copy; and

a copyright infringement information generator operative to generate a display of the copyright information.

13. A system according to claim 12 wherein at least a portion of said comparison is conducted at the shot level.

14. A system according to claim 12 or claim 13 wherein at least a portion of said comparison is conducted at the frame level.

15. A system according to claim 12 wherein the copyright information quantifies the infringement of copyright of the original image sequence by the suspected pirate copy.

16. A watermarking method comprising:

providing an image sequence to be watermarked; and

performing a predetermined alteration of the length of the image sequence.

17. A method according to claim 16 wherein said performing step comprises duplicating at least one predetermined image in the image sequence.

18. A method according to claim 16 wherein said performing step comprises omitting at least one predeter-

mined image from the image sequence.

19. A system according to claim 7 wherein the image sequence analyzer is operative to generate aligned representations of the first and second image sequences and the display unit is operative to display the aligned representations on a single screen.

20. A video sequence viewing method comprising:

displaying a sequence of images at a speed determined in accordance with a control signal; and

performing an analysis of the sequence of images and generating the control signal in accordance with a result of the analysis.

21. An image sequence viewing method comprising:

performing an analysis of a sequence of images and to identify shots within the sequence of images; and

sequentially displaying at least one initial images of each identified shot.

22. A method for displaying a first image sequence as aligned relative to a second, related image sequence, the method comprising:

generating a representation of a first image sequence including at least one row of pixels of each image in the first image sequence; and

displaying the rows generated by the analyzer, side by side, in a single screen, wherein gaps are provided between the rows, in order to denote images which are missing, relative to the second image sequence.

23. A copyright monitoring method comprising:

conducting a comparison between an original image sequence and a suspected pirate copy of the original image sequence and to generate copyright information

describing infringement of copyright of the original image sequence by the suspected pirate copy; and

generating a display of the copyright information.

24. A watermarking system comprising:

an image sequence input device operative to input an image sequence to be watermarked; and

an image sequence length alteration device operative to perform a predetermined alteration of the length of the image sequence.

25. A DVD authoring method comprising:

performing a DVD authoring operation on a plurality of versions of a motion picture, the performing step comprising:

synchronizing the plurality of versions of the motion picture, including:

capturing at least one signatures of at least one corresponding video frames within the plurality of versions of the motion pictures, using only small amounts of data to characterize each of said video frames; and

matching said signatures to a continuous stream of data.

26. An advertisement verification method comprising:

comparing a broadcast of a commercial with an original commercial, at least partly on the frame level, including comparing individual frames of the broadcast to individual frames of the original commercial; and

generating an output indicating at least one parameter of similarity between the broadcast and the original commercial.

27. A method according to claim 26 wherein the comparing step comprises at least one of the following steps:

signature extraction; and

signature search.

28. A DVD authoring method comprising:

generating a generic version of a motion picture by comparing and combining a plurality of original video clips representing said motion picture, at the frame level; and

creating branching instructions for playback of at least one subsequence of the generic version on a DVD player.

29. A DVD authoring method comprising:

creating branching instructions for playback of at least one subsequence of a generic version of a motion picture on a DVD player, the generic version comprising a combination of a plurality of original video clips representing said motion picture; and

employing said branching instructions to play back at least one subsequence and comparing said at least one subsequence, at the frame level, to at least a portion of at least one of the plurality of original video clips representing said motion picture.

30. An automated video duplication quality control method comprising:

comparing actual video content derived from a reference video content, with the reference content, thereby to obtain a measure of duplication quality control quantifying at least one aspect of similarity between the actual and reference video contents, the comparing step comprising:

extracting frame characteristic data

streams from said reference content and from said actual content;

aligning at least a portion of said streams; and

comparing at least a portion of said streams on a frame-by-frame basis.

31. A method for comparing a final DVD version of a video clip against an original clip from which the final DVD version was generated, the method comprising:

extracting characteristic data from a first audio-visual stream representing the final clip and from a second audio-visual stream representing the original clip; and

comparing the extracted characteristic data from said first and second audio-visual streams.

32. A broadcast verification system comprising:

a signature extractor operative to extract a relatively small signature from a subject clip;

a real time video scanner operative to scan a broad video stream in real time in order to identify the subject clip within the broad video stream; and

a comparison report generator operative to produce a comparison report including a frame-by-frame comparison of the subject clip and of the broad video stream.

33. A DVD authoring system comprising:

DVD authoring apparatus operative to perform DVD authoring on a plurality of versions of a motion picture, the apparatus comprising:

a synchronizer operative to synchronize the plurality of versions of the motion picture, including:

a signature capturer operative to

capture at least one signatures of at least one corresponding video frames within the plurality of versions of the motion pictures, using only small amounts of data to characterize each of said video frames; and

a signature matcher operative to match said signatures to a continuous stream of data.

34. An advertisement verification system comprising:

frame level broadcast evaluation apparatus operative to compare a broadcast of a commercial with an original commercial, at least partly on the frame level, including comparing individual frames of the broadcast to individual frames of the original commercial; and

a similarity output generator operative to generate an output indicating at least one parameter of similarity between the broadcast and the original commercial.

35. A DVD authoring system comprising:

a generic version generator operative to generate a generic version of a motion picture by comparing and combining a plurality of original video clips representing said motion picture, at the frame level; and

a brancher operative to create branching instructions for playback of at least one subsequence of the generic version on a DVD player.

36. A DVD authoring system comprising:

a brancher operative to create branching instructions for playback of at least one subsequence of a generic version of a motion picture on a DVD player, the generic version comprising a combination of a plurality of original video clips representing said motion picture; and

a frame level playback evaluator operative to

employ said branching instructions to play back at least one subsequence and comparing said at least one subsequence, at the frame level, to at least a portion of at least one of the plurality of original video clips representing said motion picture.

37. An automated video duplication quality control system comprising:

a duplication quality controller operative to compare actual video content derived from a reference video content, with the reference content, thereby to obtain a measure of duplication quality control quantifying at least one aspect of similarity between the actual and reference video contents, the controller comprising:

a frame characteristic extractor operative to extract frame characteristic data streams from said reference content and from said actual content;

a stream aligner operative to align at least a portion of said streams; and

stream comparing apparatus operative to compare at least a portion of said streams on a frame-by-frame basis.

38. A system for comparing a final DVD version of a video clip against an original clip from which the final DVD version was generated, the system comprising:

a characteristic data extractor operative to extract characteristic data from a first audio-visual stream representing the final clip and from a second audio-visual stream representing the original clip; and

apparatus for comparing the extracted characteristic data from said first and second audio-visual streams.

39. A broadcast verification method comprising:

extracting a relatively small signature from a

subject stream of video frames; and

producing a comparison report including a frame-by-frame comparison of the subject stream and of an additional video stream based on a signature-level match between the two streams.

40. A broadcast verification method comprising:

comparing a broadcast video sequence with an original video sequence, at least partly on the frame level, including comparing at least a derivation of individual frames of the broadcast to at least a derivation of individual frames of the original video sequence; and

generating an output indicating at least one parameter of similarity between the broadcast and the original video sequence.

41. A method according to claim 40 wherein said derivation of a first individual frame which is compared to a derivation of a second individual frame, in the course of said comparing step, comprises a signature of the first individual frame.

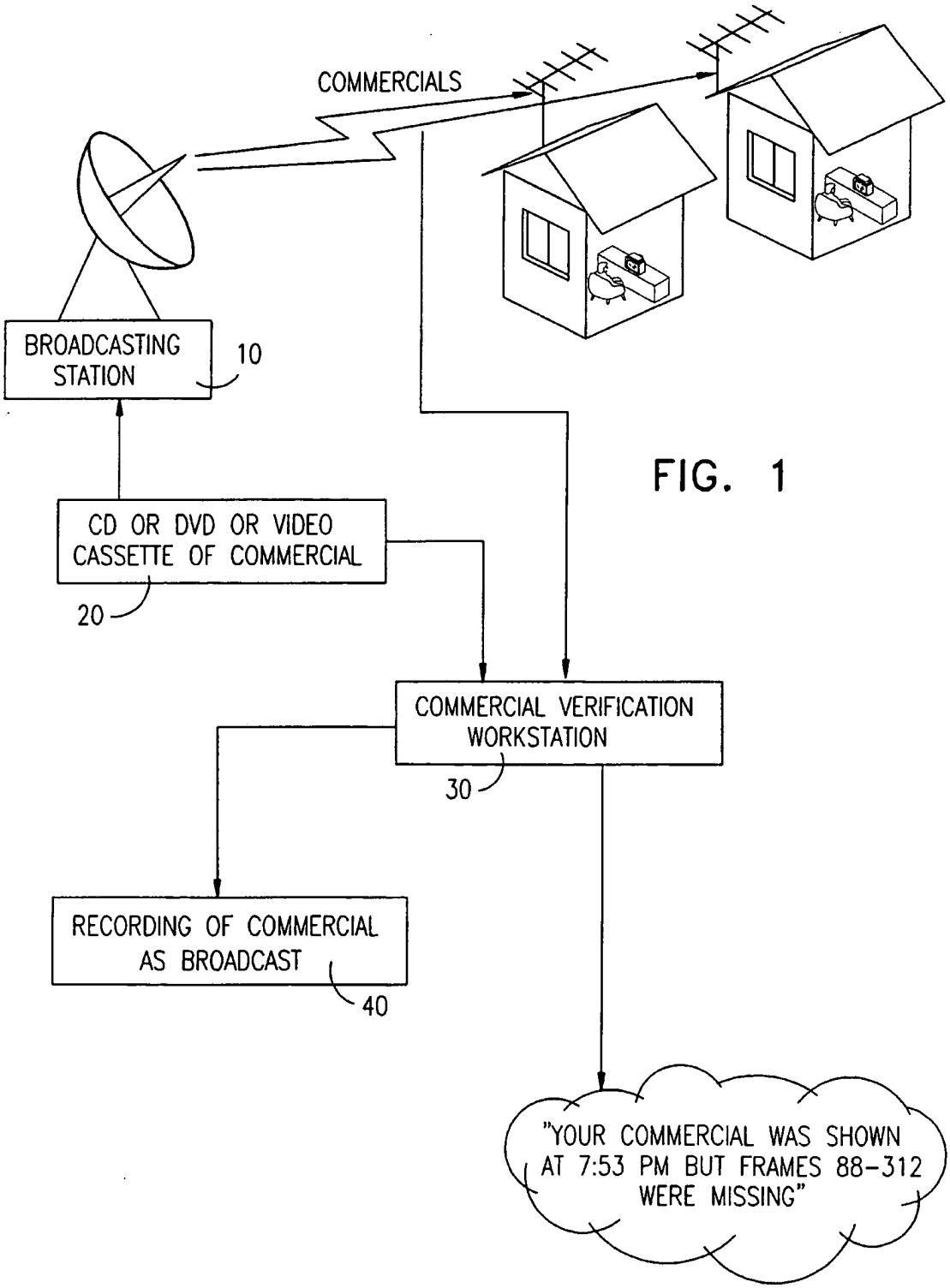

COMMERCIALS
BROADCASTING STATION
10
CD OR DVD OR VIDEO CASSETTE OF COMMERCIAL
20
COMMERCIAL VERIFICATION WORKSTATION
30
RECORDING OF COMMERCIAL AS BROADCAST
40
"YOUR COMMERCIAL WAS SHOWN AT 7:53 PM BUT FRAMES 88-312 WERE MISSING"

FIG. 1

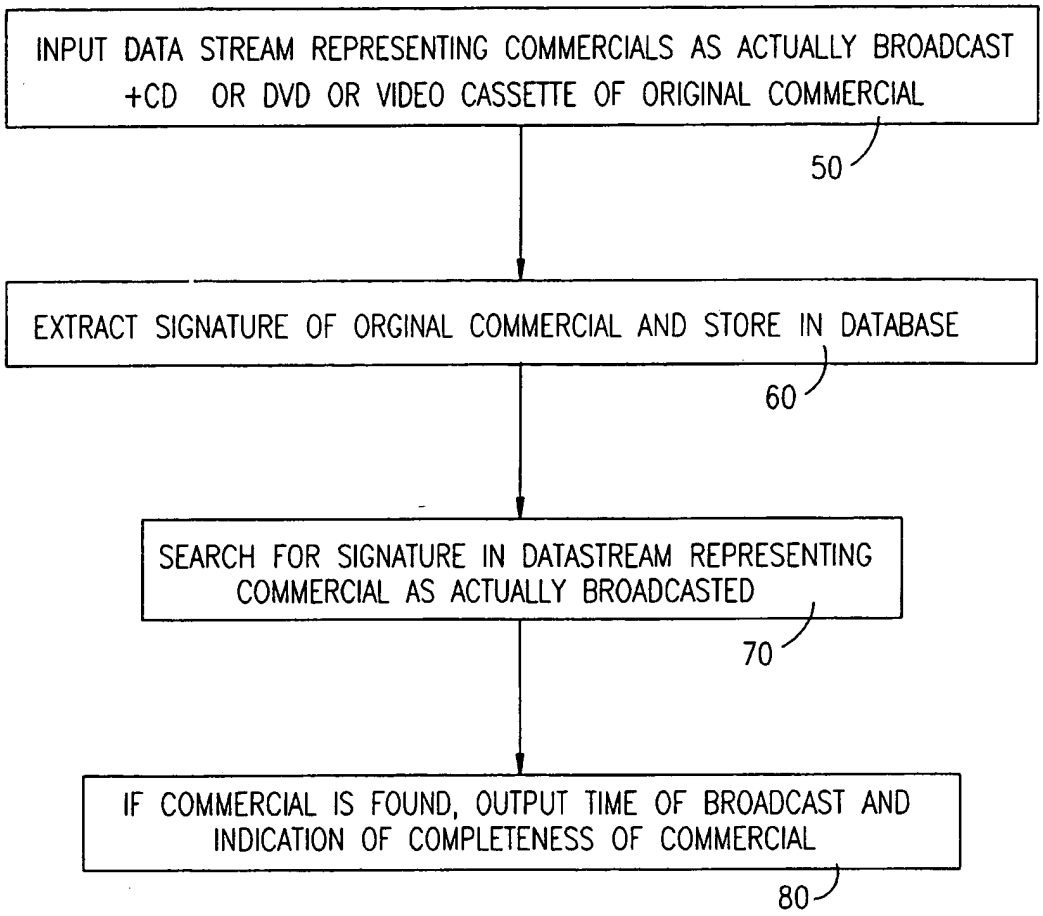

INPUT DATA STREAM REPRESENTING COMMERCIALS AS ACTUALLY BROADCAST
+CD OR DVD OR VIDEO CASSETTE OF ORIGINAL COMMERCIAL
50
EXTRACT SIGNATURE OF ORGINAL COMMERCIAL AND STORE IN DATABASE
60
SEARCH FOR SIGNATURE IN DATASTREAM REPRESENTING
COMMERCIAL AS ACTUALLY BROADCASTED
70
IF COMMERCIAL IS FOUND, OUTPUT TIME OF BROADCAST AND
INDICATION OF COMPLETENESS OF COMMERCIAL
80

FIG. 2

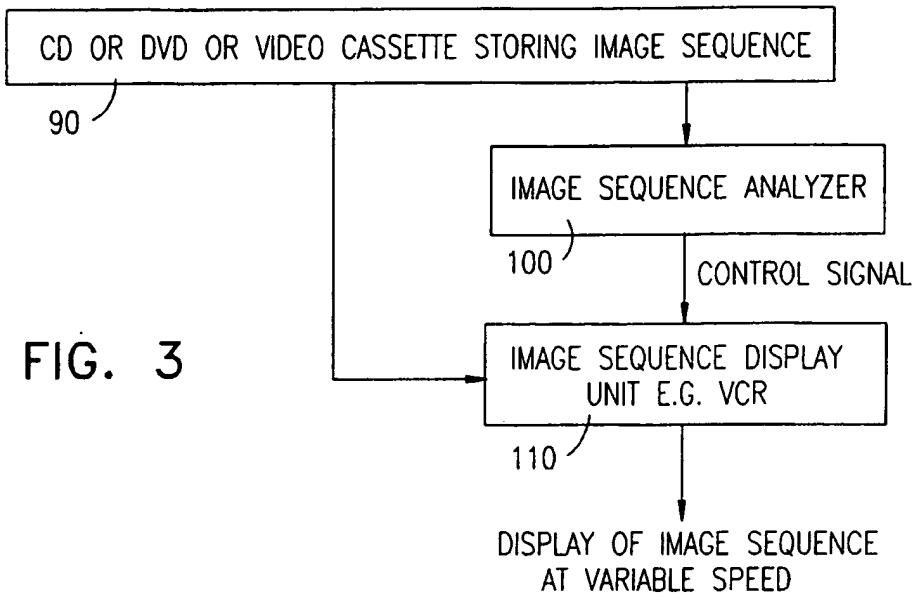

CD OR DVD OR VIDEO CASSETTE STORING IMAGE SEQUENCE
90
IMAGE SEQUENCE ANALYZER
100
CONTROL SIGNAL
IMAGE SEQUENCE DISPLAY UNIT E.G. VCR
110
DISPLAY OF IMAGE SEQUENCE AT VARIABLE SPEED

FIG. 3

FIG. 4

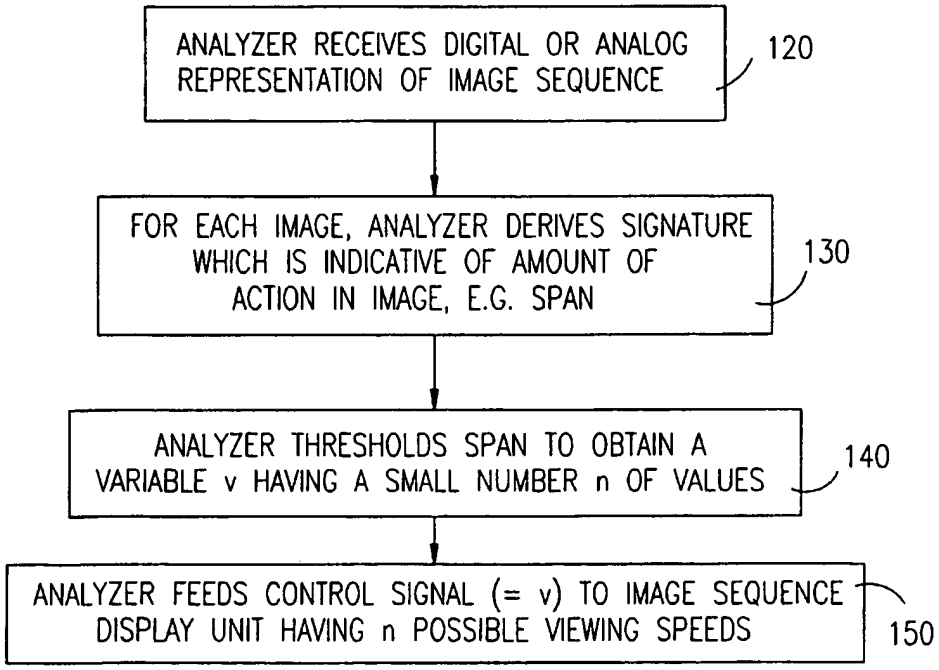

ANALYZER RECEIVES DIGITAL OR ANALOG REPRESENTATION OF IMAGE SEQUENCE
120
FOR EACH IMAGE, ANALYZER DERIVES SIGNATURE WHICH IS INDICATIVE OF AMOUNT OF ACTION IN IMAGE, E.G. SPAN
130
ANALYZER THRESHOLDS SPAN TO OBTAIN A VARIABLE v HAVING A SMALL NUMBER n OF VALUES
140
ANALYZER FEEDS CONTROL SIGNAL (= v) TO IMAGE SEQUENCE DISPLAY UNIT HAVING n POSSIBLE VIEWING SPEEDS
150

FIG. 5

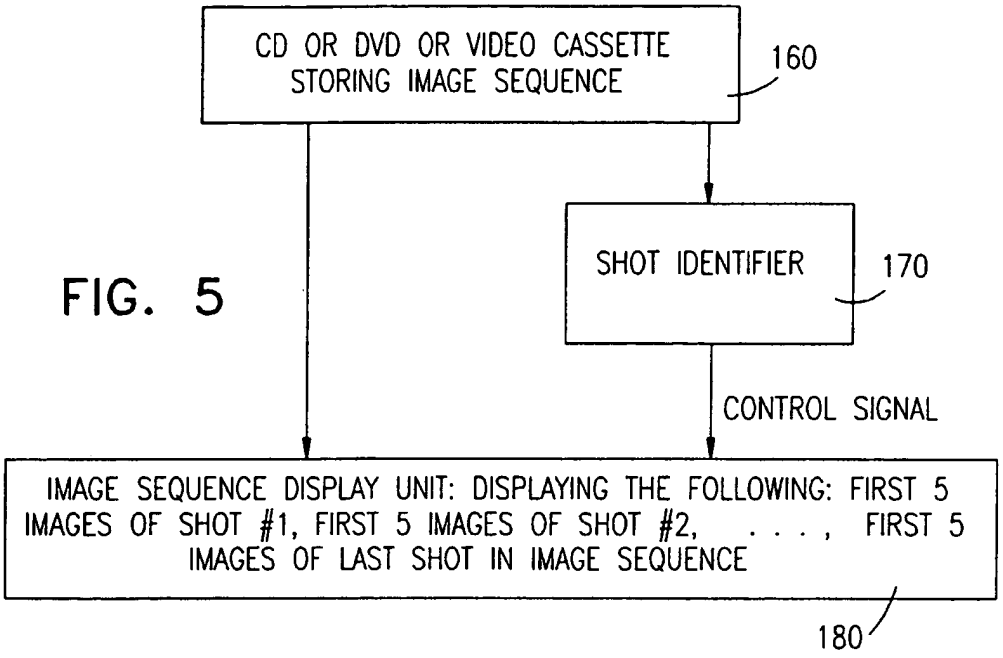

CD OR DVD OR VIDEO CASSETTE STORING IMAGE SEQUENCE
160
SHOT IDENTIFIER
170
CONTROL SIGNAL
IMAGE SEQUENCE DISPLAY UNIT: DISPLAYING THE FOLLOWING: FIRST 5 IMAGES OF SHOT #1, FIRST 5 IMAGES OF SHOT #2, . . . , FIRST 5 IMAGES OF LAST SHOT IN IMAGE SEQUENCE
180

FIG. 6

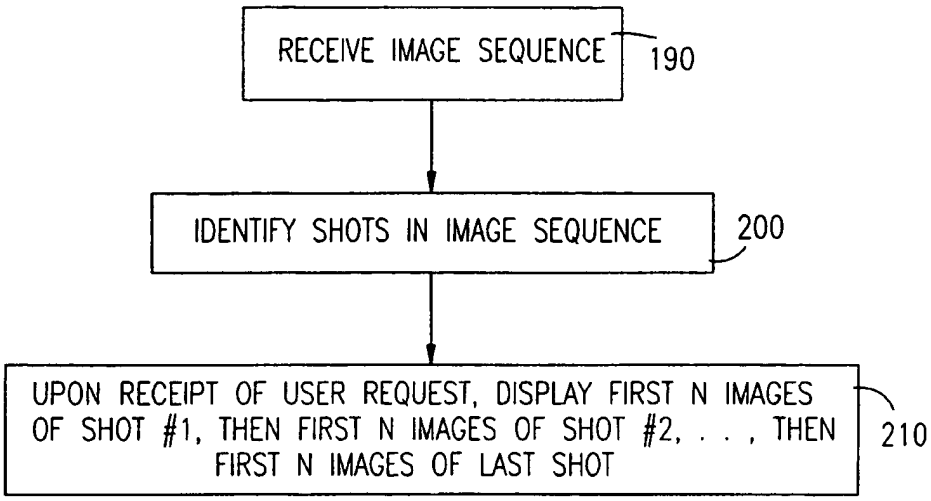
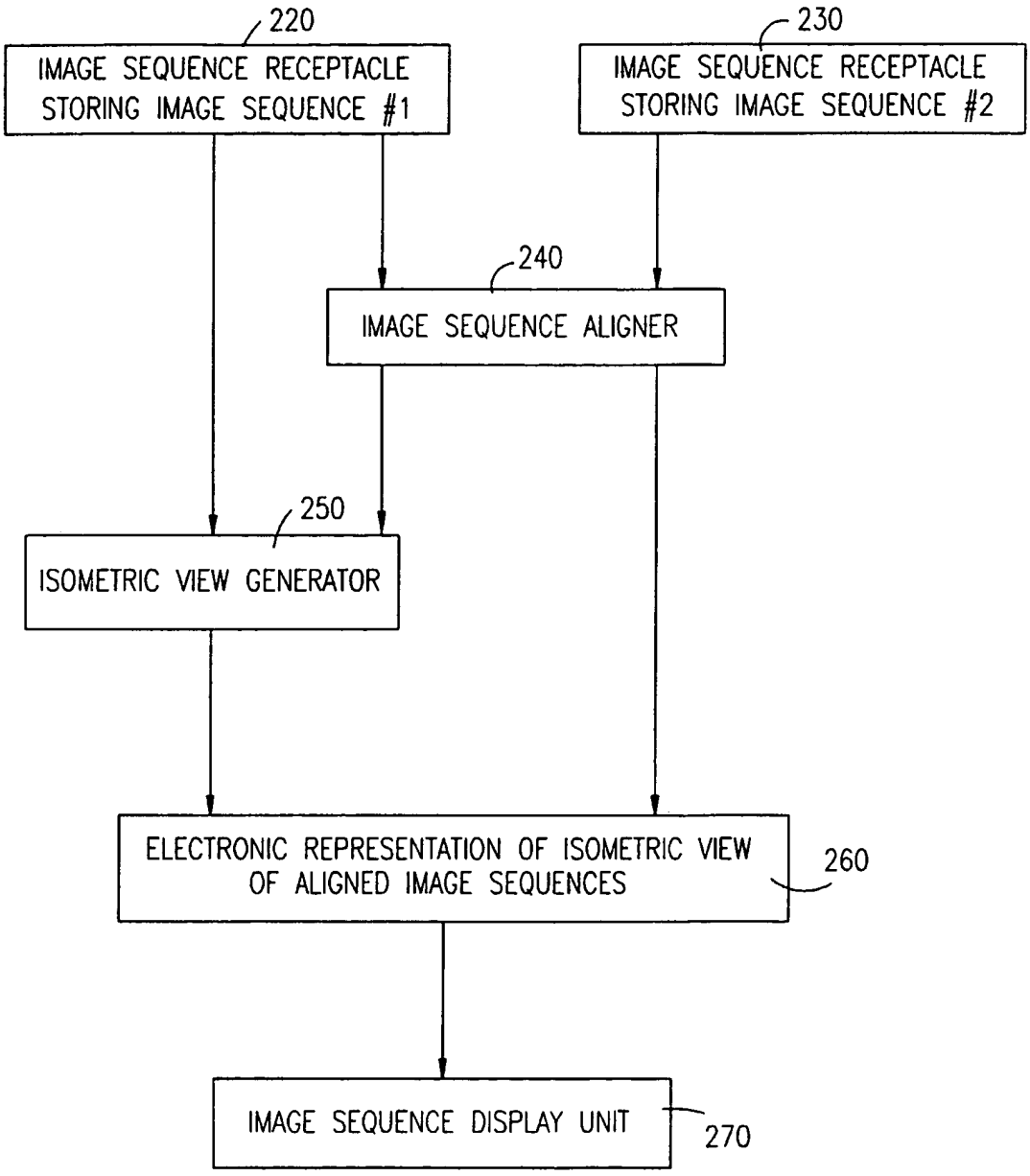

RECEIVE IMAGE SEQUENCE
190
IDENTIFY SHOTS IN IMAGE SEQUENCE
200
UPON RECEIPT OF USER REQUEST, DISPLAY FIRST N IMAGES OF SHOT #1, THEN FIRST N IMAGES OF SHOT #2, . . , THEN FIRST N IMAGES OF LAST SHOT
210

FIG. 7


220
IMAGE SEQUENCE RECEPTACLE STORING IMAGE SEQUENCE #1
230
IMAGE SEQUENCE RECEPTACLE STORING IMAGE SEQUENCE #2
240
IMAGE SEQUENCE ALIGNER
250
ISOMETRIC VIEW GENERATOR
ELECTRONIC REPRESENTATION OF ISOMETRIC VIEW OF ALIGNED IMAGE SEQUENCES
260
IMAGE SEQUENCE DISPLAY UNIT
270

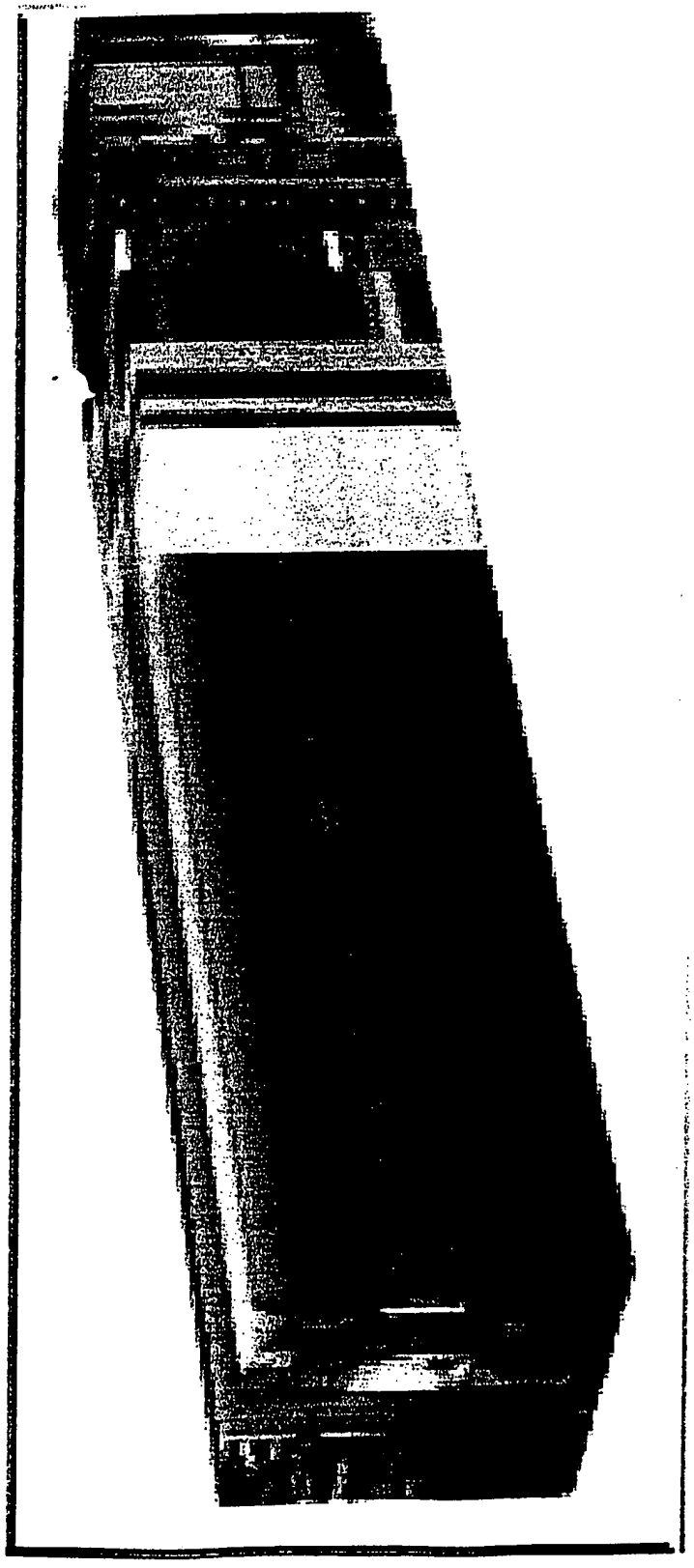


Fig. 8

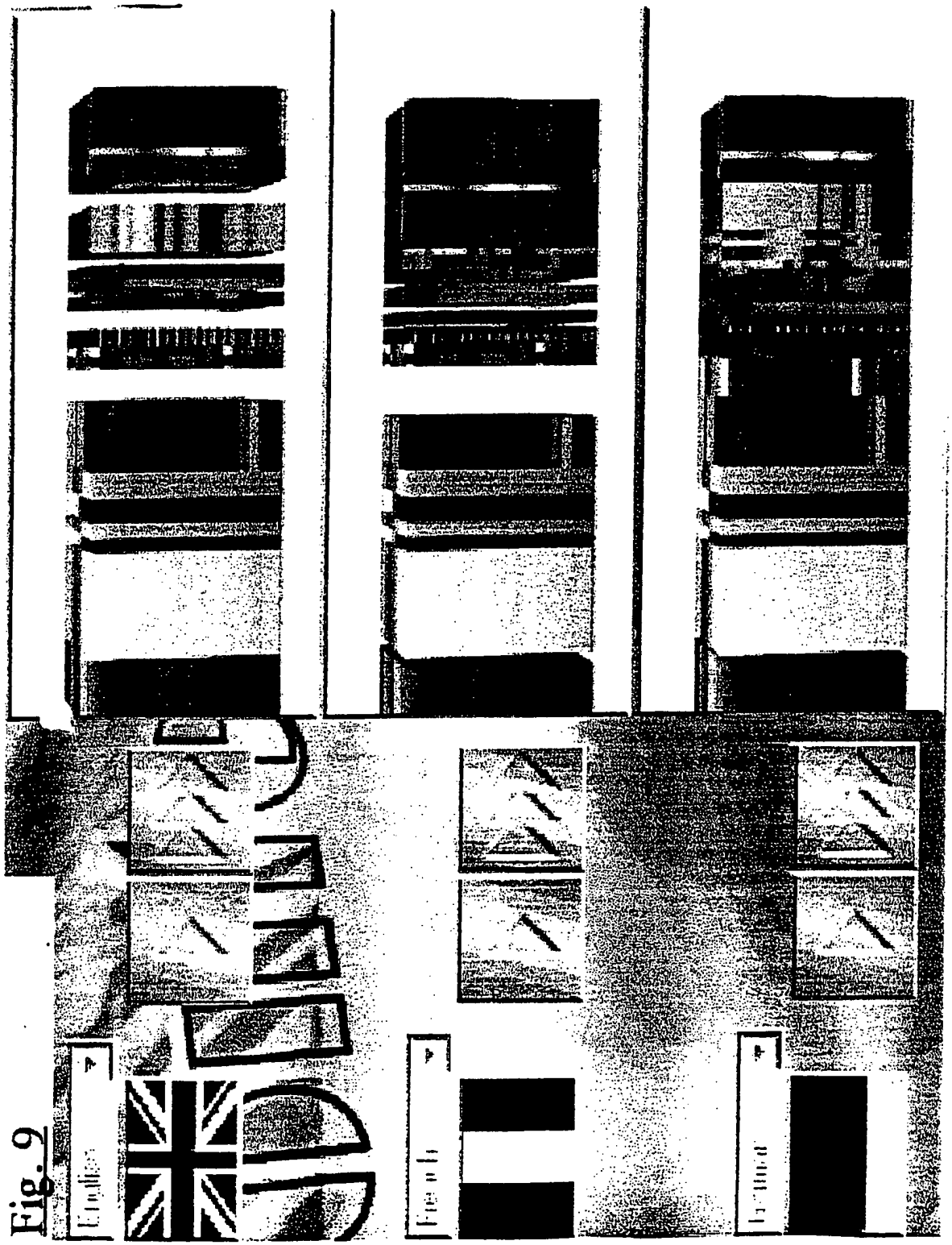


Fig. 9

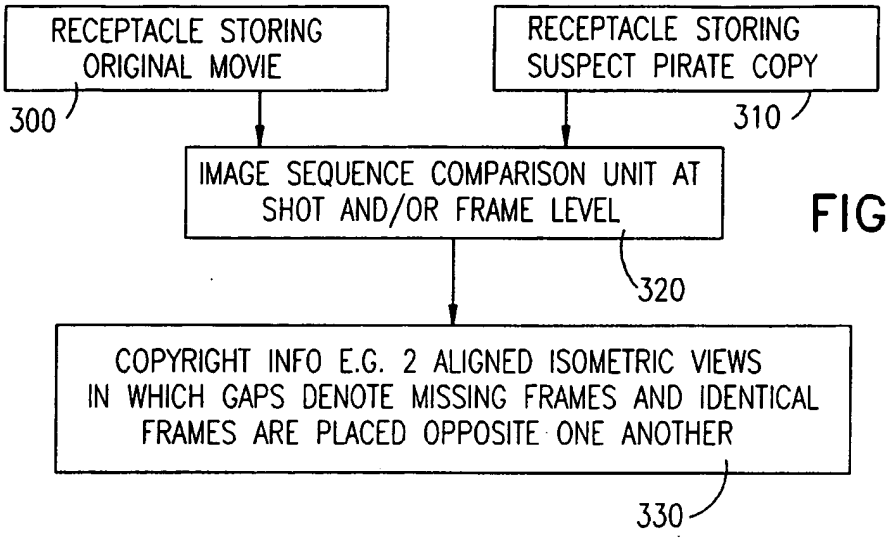
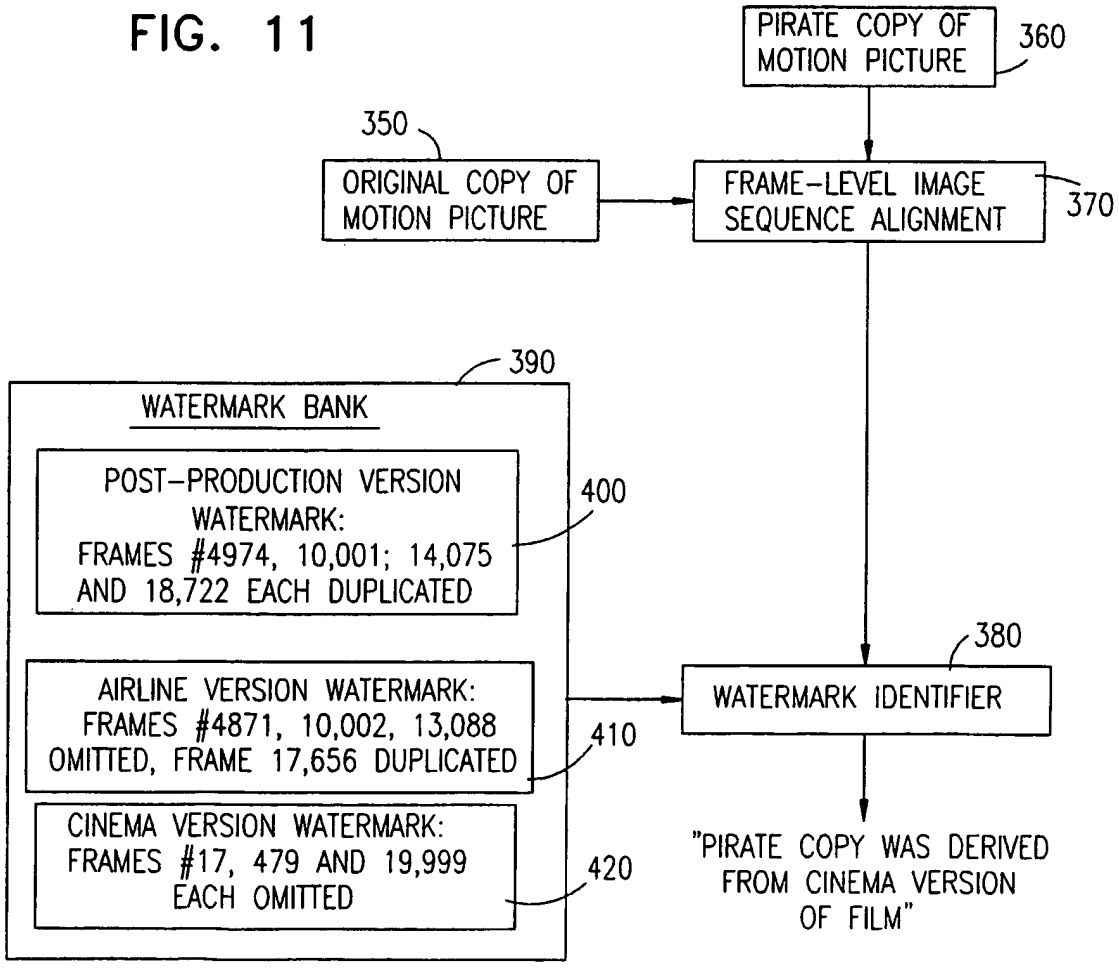

RECEPTACLE STORING ORIGINAL MOVIE
300
RECEPTACLE STORING SUSPECT PIRATE COPY
310
IMAGE SEQUENCE COMPARISON UNIT AT SHOT AND/OR FRAME LEVEL
320
COPYRIGHT INFO E.G. 2 ALIGNED ISOMETRIC VIEWS IN WHICH GAPS DENOTE MISSING FRAMES AND IDENTICAL FRAMES ARE PLACED OPPOSITE ONE ANOTHER
330

FIG. 10

FIG. 11


PIRATE COPY OF MOTION PICTURE
360
350
ORIGINAL COPY OF MOTION PICTURE
FRAME-LEVEL IMAGE SEQUENCE ALIGNMENT
370
390
WATERMARK BANK
POST-PRODUCTION VERSION WATERMARK: FRAMES #4974, 10,001; 14,075 AND 18,722 EACH DUPLICATED
400
AIRLINE VERSION WATERMARK: FRAMES #4871, 10,002, 13,088 OMITTED, FRAME 17,656 DUPLICATED
410
CINEMA VERSION WATERMARK: FRAMES #17, 479 AND 19,999 EACH OMITTED
420
380
WATERMARK IDENTIFIER
"PIRATE COPY WAS DERIVED FROM CINEMA VERSION OF FILM"

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL98/00596

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :H04N 5/222
US CL :348/722

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 348/722, 1, 4, 5, 459, 500, 575, 576, 700, 715, 723, 907; 345/327

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,680,454 A (MEAD) 21 October 1997, See Whole document	1, 2, and 20
X	US 5,642,174 A (KAZUI et al) 24 June 1997, See Whole document	4
X,P	US 5,842,023 A (TSUMURA) 24 November 1998, See Abstract	12
X	US 5,848,155 A (COX) 08 December 1998, See Whole document	16
X	US 5,659,613 A (COPELAND et al) 19 August 1997, See Abstract	25
X	US 5,504,518 A (ELLIS et al) 02 April 1996, See Abstract	26, 40, and 41

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
22 MARCH 1999	16 APR. 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	HABTE K. BAHGI
Facsimile No. (703) 305-3230	Telephone No. (703) 308-8208

Form PCT/ISA/210 (second sheet)(July 1992) *

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL98/00596

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,870,754 A (DIMITROVA et al) 09 February 1999, See Whole document	30-32
X	US 5,784,464 A (AKIYAMA et al) 21 July 1998, See Whole document	33-38
A	US 5,646,675 A (COPRIVIZA et al) 08 July 1997, See Whole document	26, 32, 34, 39, and 40
A	US 5,621,454 A (ELLIS et al) 15 April 1997, See Whole document	26, 32, 34, 39, and 40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IL98/00596

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS
Search terms: video image?, image sequence analizer, shot identif?, copyright monitor#, information generator, watermark?, authoring, dvd authoring, motion picture, video frame, signature extract?, signature search compar? broadcast, advertise? verification, automated video duplication

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.